

جامعة مؤتة

كلية الدراسات العليا

استخدام البواقي والقيم الشاذة للكشف عن انتهاكات افتراضات تحليل الانحدار الخطي البسيط

إعداد الطالب

جميل فرهود أبو قديري

إشراف

الدكتور صبري حسن الطراونة

رســـــالة مقدمـــــة إلــــى كليـــــة الدراســــــات العليـــــا

استكمالاً لمتطلبات الحصول على درجة الماجستير في

القيـــــــاس والتقـــــويم/ قســــــم علـــــــم الـــــــنفس

جامعة مؤتة 2016

الآراء الواردة في الرسالة الجامعية لا تعبر بالضرورة عن وجهة نظر جامعة مؤتة

الإهداء

لو كان يُهدى للإنسان قيمته لكنت أهديت الدنيا وما فيها

إلى الأرواح الغائبة الحاضرة أمي وأبي.

إلى من عشت معهم أجمل السنين شقيقاتي وأشقائي.

إلى رياحين بستاني نتاشا ، جمان ، كرم ، بهاء.

إلى رفيقة دربي زوجتي.

إلى كل من تعلمت منهم حرفاً مدرسين وزملاء

أُهدي هذا العمل الذي أسال الله أن يُنتفع منه.

جميل فرهود أبو قديري

الشكر والتقدير

الحمد لله رب العالمين والصلاة والسلام على الأنبياء أجمعين بدءاً بآدم وانتهاءً بالنبي الأمين سيدنا محمد عليه أفضل الصلوات والتسليم.

لا يسعني إلا الوقوف والتقدم بوافر الشكر والتقدير إلى أخي وأستاذي الفاضل الدكتور صبري الطراونة على دعمه النفسي والعلمي والجهود الكبيرة التي بذلها معي منذ أن خط قلمي أول كلمات هذه الدراسة وكان لجهوده الأثر الكبير في إخراج هذه الرسالة بالشكل الذي هي عليه، وأسال الله أن تكون جهودكم في موازين حسناتكم، والشكر موصول إلى جميع أعضاء الهيئة التدريسية في كلية العلوم التربوية وزملائي الطلبة.

وأتقدم بجزيل الشكر إلى أعضاء لجنة المناقشة الذين تجشموا عناء قراءة الرسالة وتقويم اعوجاجها وسد خللها.

جميل فرهود أبو قديري

فهرس المحتويات

الصفحة	المحتوى

قائمة الجداول

قائمة الأشكال

قائمة الملاحق

الملخص

استخدام البواقي والقيم الشاذة للكشف عن انتهاكات افتراضات تحليل الانحدار الخطي البسيط

جميل فرهود أبو قديري

جامعة مؤته، 2016

هدفت الدراسة إلى الكشف عن انتهاكات افتراضات تحليل الانحدار الخطي البسيط ودقة معادلة الانحدار، وذلك من خلال تناول البواقي والقيم الشاذة، وتم استخدام مجموعة من الاختبارات (نقص المطابقة، تجانس التباين، واستقلالية البواقي، والتوزيع الطبيعي) للكشف عن انتهاك افتراضات تحليل الانحدار، وتبين من خلال الاختبارات أن القيم الكبيرة لمعامل التحديد وكذلك معنوية اختبار (F) ومعنوية معاملات معادلة الانحدار ليست دليلاً كافياً على أن البيانات مطابقة لنموذج الانحدار الخطي، ولا يمكن الاعتماد عليها وحدها للحكم على دقة النموذج الخطي. حيث أظهرت نتائج اختبار نقص المطابقة للبيانات التي تحتوي على قيم مكررة في المتغير المستقل أن بعض البيانات لا تلائم النموذج الخطي. كما أظهرت نتائج اختبار تجانس التباين عدم تحقق افتراض تجانس التباين في بعض البيانات. وكذلك أظهرت النتائج انتهاك افتراضات الاستقلالية والتوزيع الطبيعي في بعض البيانات، وأظهرت النتائج أن معالجة القيم الشاذة يعمل على تحقيق افتراض استقلالية البواقي وافتراض التوزيع الطبيعي، وكما أظهرت نتائج الدراسة أن وجود القيم الشاذة يؤدي إلى عدم تحقق افتراض تجانس التباين للبواقي وافتراض نقص المطابقة.

Abstract

Using Residuals and Outliers for Detecting Violations of Simple Linear Regression Analysis Assumptions.

Jamil farhood Abu qudeiri

Mu'tah University, 2016

The study aimed to detect violations simple linear regression and accuracy regression analysis assumptions by using residuals and outlier values . A set of tests were carried out in this research such as (Lack of Fit, Homoscedasticity , Independence of Residuals, Normal Distribution) in order to detect violation of assumptions. The results of tests showed that the large values for coefficient of determination as well as the significance of (F) test and significance of regression equation coefficients are not sufficient evidence that data is identical to the linear regression model and it can't be relied on alone to judge the accuracy of linear model.

Lack of Fit test results for data which contains duplicated values of independent variable showed that some data doesn't fit the linear model. Homoscedasticity test results indicated that homoscedasticity assumption hasn't been achieved in some data. As well as the results showed violation of Independence of Residuals and Normal Distribution assumptions in some data. The research has tackled outliers and its treatment ,and highlighted its role in violation of assumptions and accuracy of linear regression equation. The findings showed that outliers treatment led to achieve Independence of Residuals and Normal Distribution assumptions . The study also showed that the presence of outliers was the cause of Heteroscedasticity and Lack of Fit assumptions.

الفصل الأول

خلفية الدراسة وأهميتها

1.1 مقدمة

يهدف العلم إلى تحقيق أربعة أهداف رئيسة متمثلة في الوصف، والتفسير، والتنبؤ والضبط والتحكم، ولما كان الهدف من تحليل الانحدار هو إيجاد دالة العلاقة بين المتغيرات المستقلة والمتغير التابع والتي تساعد في تفسير التغير الذي قد يطرأ على المتغير التابع تبعاً للتغير في قيم المتغيرات المستقلة، والتنبؤ بقيمة المتغير التابع لقيمة محددة من المتغير المستقل؛ أي إن تحليل الانحدار يسعى لتحقيق أهداف العلم- فلا بد من اختيار نوع الانحدار الذي يلائم البيانات المدروسة لتحقيق الهدف الذي جاء من أجله.

ويعد تحليل الانحدار الخطي البسيط طريقة إحصائية تهدف إلى تحديد العلاقة الرياضية بين المتغير التابع والمتغير المستقل من خلال رسم أفضل خط مستقيم يلائم البيانات، والذي يمتلك أصغر مجموع لمربعات المسافة بين القيم المشاهدة والقيم المتوقعة للمتغير التابع باستخدام طريقة المربعات الصغرى، التي يتم من خلالها حساب معلمات معادلة الانحدار الخطي البسيط:

$$Y_i = \beta_0 + \beta_{1xi} + \varepsilon_i \qquad (1)$$

حيث (y_i) المتغير التابع و(xi) المتغير المستقل وقيمة (ε_i) هو الخطأ العشوائي (البواقي) و(β_0) قيمة التقاطع وهي القيمة المتوقعة للمتغير التابع عندما تكون قيمة المتغير المستقل تساوي صفرا، وقيمة (β_1) تعبر عن ميل الخط المستقيم وهي التغير في قيمة المتغير التابع عندما تتغير قيمة المتغير المستقل وحدة واحدة (نوروسيس ، 2010).

ومن المهم أن نذكر هنا أنه من بين الأخطاء الشائعة في تحليل نماذج الانحدار اعتماد نتيجة جدول تحليل التباين للانحدار ومعنوية معلماته بدون معالجة مجموع مربعات الخطأ الكلي وتأثيره على قيمة الاختبار (F) الأمر الذي يترتب عليه اعتماد نتائج مضللة، حيث إن القيم الكبيرة لمعامل التحديد أو معنوية اختبار (t) لا تؤكد أن

البيانات قد تم مطابقتها للنموذج بصورة جيدة (يونس وابراهيم، 2004).

وإن هذه الطريقة تكون مضللة وغير دقيقة في حالة عدم تحقق احد افتراضاتها الأساسية أو بسبب وجود القيم الشاذة التي تؤثر على معادلة الانحدار وتجعل الاستنتاجات المبنية عليها غير دقيقة، لهذا يجب في البداية التأكد من خلو البيانات من القيم الشاذة من خلال الكشف عنها ومعالجتها بالطرق التي تجعلنا واثقين من نتائج التحليل (الزعبي والطلافحه، 2012)

كما تعمل القيم الشاذة على تضخم تباين البواقي الناتجة بعد ملاءمة خط الانحدار وتؤدي إلى انتهاكات افتراضاته؛ لذا يجب تفحص البواقي والإحصائيات المرتبطة بها لمعرفة ما إذا كانت تنتهك أي من افتراضات تحليل الانحدار ومدى ملائمة خط الانحدار للبيانات من خلال إجراء عدد من الاختبارات الاحصائية(يوسف، 2015)

2.1 مشكلة الدراسة:

تتمثل مشكلة الدراسة في أن الكثير من الدارسين والباحثين المستخدمين لتحليل الانحدار يعتمدون على معنوية معالم معادلة الانحدار من خلال قيمة (F) المحسوبة في جدول تحليل التباين ،وكذلك قيمة مربع معامل الارتباط (R^2)، ومعنوية كل من قيمتي الميل والثابت لمعادلة الانحدار دون التأكد من تحقيق جميع افتراضات تحليل الانحدار الأساسية والضرورية للحصول على النموذج الدقيق.

ويعد التحقق من افتراضات العلاقة الخطية والتوزيع الطبيعي والتباين الثابت للأخطاء العشوائية واستقلالية المشاهدات جزءاً هاما من تحليل الانحدار، ومخالفة أحد هذه الافتراضات يؤدي إلى أخطاء في عمليتي التنبؤ والتقدير، وعدم دقة القرارات المبنية على نتائج التحليل، وقد يُعزى السبب في مخالفة أحد هذه الافتراضات إلى القيم الشاذة المؤثرة في بيانات تحليل الانحدار مما يؤدي إلى تضخم تباين البواقي وهذا يقلل من تقدير معالم النموذج، أو إلى عدم استخدام نوع تحليل الانحدار الملائم للبيانات (الراوي، 1987). ولهذا جاءت الحاجة إلى هذه الدراسة التي ستستخدم القيم الشاذة والبواقي للكشف عن مدى تحقق افتراضات نموذج تحليل الانحدار البسيط ودقة القرارات المبنية عليها، وبالتحديد فإن هذه الدراسة حاولت الإجابة عن الأسئلة التالية:

1-هل تعد معنوية نموذج الانحدار الخطي البسيط ومعاملاته دليلا كافيا على تحقق افتراضاته؟

2- ما أثر القيم الشاذة ومعالجتها على دقة معادلة الانحدار الخطي البسيط ؟

3- ما أثر القيم الشاذة ومعالجتها على افتراضات بواقي تحليل الانحدار ؟

3.1 أهداف الدراسة:

هدفت الدراسة إلى:

1- الكشف عن انتهاكات افتراضات تحليل الانحدار الخطي.

2- دراسة تأثير القيم الشاذة على افتراضات تحليل الانحدار الخطي البسيط .

3-تحليل بواقي الانحدار للتحقق من افتراضات تحليل الانحدار.

4- دارسة تأثير وجود ومعالجة القيم الشاذة في البيانات على معالم نموذج تحليل الانحدار الخطي البسيط.

4.1 أهمية الدراسة:

تتمثل أهمية الدراسة في استخدام القيم الشاذة وتحليل بواقي الانحدار الخطي البسيط للكشف عن انتهاكات افتراضات تحليل الانحدار الخطي البسيط ومدى تأثير القيم الشاذة على معالم معادلة الانحدار الخطي وانتهاك افتراضاته الأساسية . وللدراسة أهمية أخرى تتمثل في طرق الكشف عن القيم الشاذة باستخدام الرسم الصندوقي وتحليل البواقي وطرق معالجة القيم الشاذة للتقليل من أثرها على معالم معادلة الانحدار وتحقيق افتراضات تحليل الانحدار وبالتالي الوصول إلى معادلة موثوق بها .

5.1 حدود الدراسة:

استخدمت الدراسة بيانات تحليل الانحدار أخذت من بيانات حقيقية واردة في المراجع المتخصصة والدراسات السابقة التي تناولت موضوع تحليل الانحدار الخطي البسيط.

6.1 مفاهيم الدراسة:

القيم الشاذة (Outliers Values) : هي القيم التي تقع بعيداً عن خط الانحدار ويكون حد الخطأ لها كبيراً مقارنة ببقية القيم الأخرى.

البواقي (Residual) : هي ما يتبقى بعد ملاءمة خط الانحدار وتقدر قيمتها بحساب الفرق بين القيمة المشاهدة للمتغير التابع والقيمة المتوقعة له، وهذه القيمة تكون موجبة للقيم الواقعة أعلى خط الانحدار وسالبة للقيم الواقعة أسفله.

الانحدار الخطي البسيط (Simple linear regression) : هو نموذج إحصائي يقوم بتقدير العلاقة التي تربط متغير كمي واحد وهو المتغير التابع مع متغير كمي آخر وهو المتغير المستقل، وينتج من هذا النموذج معادلة احصائية خطية يمكن استخدامها لتفسير العلاقة بين المتغيرين أو تقدير قيمة المتغير التابع عند معرفة قيمة المتغير المستقل. ويمكن صياغة العلاقة بالنموذج التالي :

$$Y_i = \beta_0 + \beta_{1xi} + \varepsilon_i \qquad (2)$$

الفصل الثاني

الإطار النظري والدراسات السابقة

2.1 الإطار النظري:

تحليل الانحدار (Regression analysis)

يهدف تحليل الانحدار إلى تقدير الصورة الرياضية للعلاقة بين المتغير التابع (Dependent variable) والمتغير أو المتغيرات المستقلة (Independent variables)، ويستخدم تحليل الانحدار لدراسة مدى تأثير متغير مستقل واحد أو أكثر على متغير تابع محدد بحيث يمكن التنبؤ بقيم المتغير التابع إذا علمت قيم المتغير المستقل أو المتغيرات المستقلة، ويجب توفر شروط أساسية لإجراء تحليل الانحدار حتى تكون النتائج دقيقة ويعُتمد عليها في عملية التنبؤ بقيم المتغير التابع (إسماعيل، 2001).

ويصنف تحليل الانحدار إلى:

1-الانحدار الخطي (Linear regression) ويقسم إلى:

أ- الانحدار الخطي البسيط (Simple linear regression) ويتكون من متغير مستقل واحد مرتبط بعلاقة خطية مع المتغير التابع.

ب- الانحدار الخطي المتعدد(Multiple regression) ويتكون من عدة متغيرات مستقلة مرتبطة بعلاقة خطية مع المتغير التابع

2-الانحدار غير الخطي(Non-linear regression) ويقسم إلى:

أ- الانحدار غير الخطي البسيط (Non-linear simple regression) ويتكون من متغير مستقل واحد مرتبط بعلاقة غير خطية مع المتغير التابع

ب- الانحدار غير الخطي المتعدد (Non-linear multiple regression) ويتكون من عدة متغيرات مستقلة مرتبطة بعلاقة غير خطية مع المتغير التابع (السيفو، شلوف وجواد، 2006)

أنواع العلاقات بين المتغيرات:

تصنف العلاقات بين متغيرات تحليل الانحدار إلى صنفين هما:

أ- علاقات تامة حتمية (Deterministic relations)

وفي هذه العلاقة لا يكون هناك مجال للخطأ في حساب قيمة المتغير التابع y إذا علمت قيمة المتغير المستقل x ويكون شكل العلاقة التي تربط بين المتغيرين كالآتي:

$$y = \alpha + \beta x \qquad (3)$$

ب- علاقات عشوائية (غير تامة) (Stochastic relations)

تنشأ العلاقة العشوائية بين متغيرين أو أكثر، إذا حصل تغير في قيمة المتغير التابع نتيجة للتغير في قيمة المتغير المستقل لأسباب خارجة عن السيطرة، وحتى يتم التنبؤ بتغير قيمة المتغير التابع (y) نتيجة التغير في قيمة المتغير المستقل (x)، لا بد من إضافة متغير عشوائي (e) ليوضح العلاقة العشوائية بين المتغيرين (y) و (x). ويسمى المتغير (e) بحد الخطأ العشوائي أو حد الخطأ أو البواقي ويكون شكل العلاقة التي تربط بين المتغيرات كالآتي:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1.........n \qquad (4)$$

حيث إن:

Y_i : المتغير التابع (dependent variable) أو المتغير المُفسَر

X_i : المتغير المستقل (independent variable) أو المتغير المفُسِر

β_0 : نقطة تقاطع خط الانحدار بالمحور y

β_1 : معامل انحدار Y_i على X_i وهو يمثل مقدار التغير في Y_i عند زيادة وحدة واحدة من X_i وإشارته تدل على اتجاه العلاقة بين المتغيرين (Y_i) و (X_i) فإذا كانت إشارته موجبة دل ذلك على أن العلاقة بين المتغيرين طردية، أي عند زيادة وحدة واحدة من X_i فإن (Y_i) تزداد بمقدار β_1. أما إذا كانت إشارته سالبة دل ذلك على أن العلاقة بين المتغيرين عكسية؛ أي عند زيادة وحدة واحدة من (X_i) فإن (Y_i) تقل بمقدار (β_1)

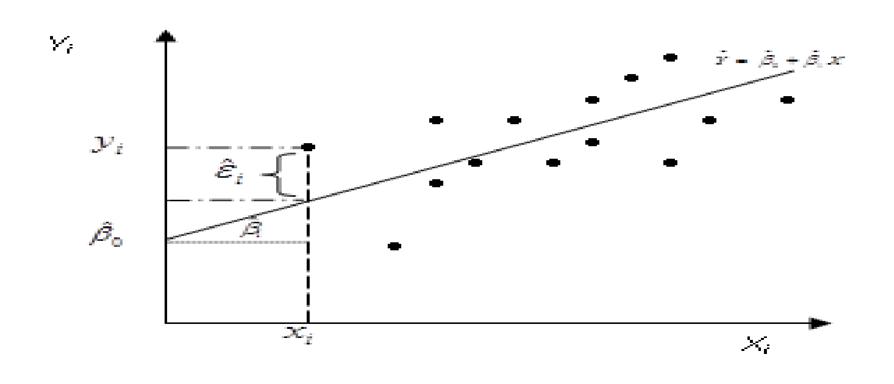
ε_i : قيمة الخطأ العشوائي (الباقي) الذي يعبر عن الفرق بين القيمة الفعلية (y)، والقيمة المقدرة $\hat{y} = \beta_0 + \beta_1 X$ (بني هاني، 2014)

إن العلاقة الواردة في المعادلة رقم (4) تمثل معادلة خط الانحدار البسيط وتوصف بأنها بسيطة؛ لأنها تحتوي على متغير مستقل واحد (X_i) وخطية لأن المعلمات β_0و β_1 خطية وأن المتغير المستقل (X_i) هو أيضاً خطي (لأنه من الدرجة الأولى) (العتوم، 2015)

تقدير نموذج الانحدار الخطي البسيط

إن الطريقة المعتمدة غالبا في تقدير معالم الانحدار β_0و β_1هي طريقة المربعات الصغرى (Least squares method) والتي باستخدامها يكون مجموع مربعات انحرافات القيم المشاهدة عن خط الانحدار المقدر أقل من مجموع مربعات انحرافات القيم عن أي خط مستقيم آخر، وتعتمد طريقة المربعات الصغرى على تقليل مجموع انحرافات القيم المشاهدة عن القيم المقدرة إلى أقل ما يمكن، وبتحقيق افتراضات نموذج تحليل الانحدار سابقة الذكر تتميز طريقة المربعات الصغرى بخصائص جيدة جعلتها من أقوى وأوسع الطرق استخداما في تقدير معالم معادلة الانحدار (إسماعيل، 2001).

والشكل (1) يبين خط الانحدار للمربعات الصغرى.



شكل (1)

خط الانحدار للمربعات الصغرى

ويتم حساب معالم معادلة الانحدار β_0و β_1 بالمعادلات الرياضية التالية:

$$\begin{cases} \hat{\beta}_1 = \dfrac{n\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{n\sum_i X_i^2 - \left(\sum_i X_i\right)^2} \\ \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \end{cases} \qquad (5)$$

ومن الممكن استخدام صيغة مكافئة لتقدير $\hat{\beta}_1$: وهي

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2} \qquad (6)$$

ويكون النموذج المقدر (خط الانحدار) بطريقة المربعات الصغرى المقدرة كما يلي: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$ (عطيه، 2000)

افتراضات الانحدار الخطي البسيط:(Simple linear regression assumption)

- أن تكون العلاقة بين المتغير التابع (y) والمتغير المستقل (x) علاقة خطية
- أن تكون كل من قياسات المتغير التابع والمتغير المستقل صحيحة.
- أن يكون تباين المتغير المستقل أكبر من الصفر.
- أن تكون القيمة المتوسطة لحد الخطأ العشوائي (البواقي) تساوي صفرا.
- أن يكون تباين حد الخطأ العشوائي (البواقي) لكل قيم المتغير المستقل ثابتا
- استقلالية قيم حدود الخطأ العشوائي (البواقي) بعضها عن بعض.
- عدم وجود ارتباط بين حد الخطأ العشوائي (البواقي) والمتغير المستقل.
- أن يتبع حد الخطأ العشوائي (البواقي) التوزيع الطبيعي (نوروسيس، 2010)

تحليل البواقي (Residuals Analysis)

عند تقدير معالم نموذج الانحدار الخطي البسيط (β_0) و(β_1) للبيانات بطريقة المربعات الصغرى الاعتيادية تتبع هذه العملية اختبارات بديهية تتمثل في اختبار(t) واختبار (F) لنموذج الانحدار وحساب قيمة معامل التحديد (R^2)، ومن بين الأخطاء الشائعة في تحليل نماذج الانحدار التي لا يقوم بها كثير من الباحثين اعتماد نتيجة جدول تحليل التباين للانحدار بدون إجراء تحليل لبواقي الانحدار، الأمر الذي يترتب عليه اعتماد نتائج مضللة، حيث إن القيم الكبيرة لمعامل التحديد أو معنوية اختبار (t)

لا تؤكد على أن البيانات قد تم مطابقتها للنموذج بصورة جيدة (يونس وإبراهيم ، 2004).

البواقي (Residuals)

وهي القيم المقدرة لحد الخطأ العشوائي(ε_i) وهي تمثل الفرق بين القيمة المشاهدة (observed value) والقيمة المتوقعة (Expected value) المناظرة لها، أي:

$$\varepsilon_i = Y_i - \beta_0 - \beta_1 X_i \qquad (7)$$

وقد تكون قيمة الباقي قيمة موجبة أو سالبة أو مساوية للصفر.

البواقي المعيارية (Standardized Residuals)

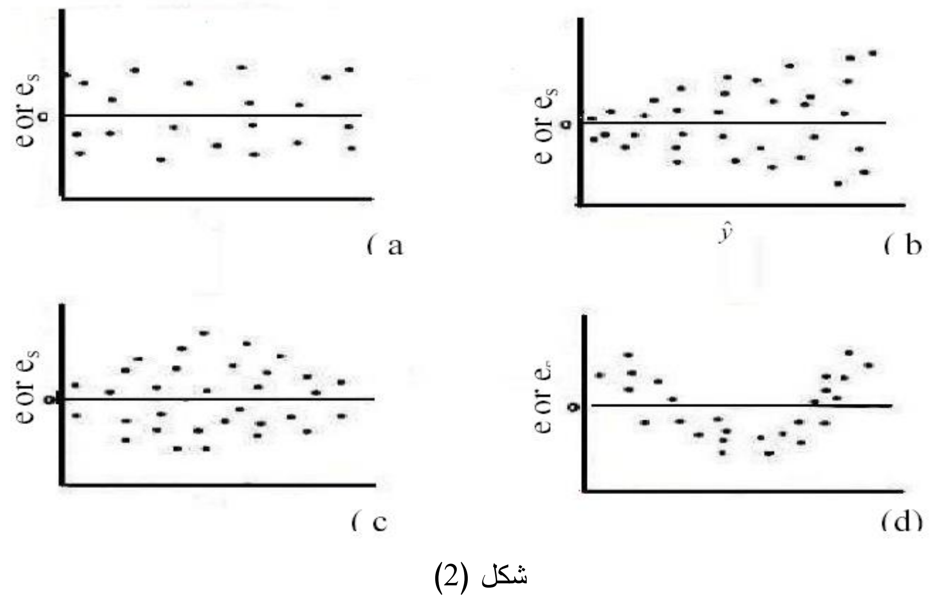
وهي القيم المقدرة لحد الخطأ العشوائي(e) مقسومة على القيمة المقدرة للانحراف المعياري للبواقي وبعد اختيار النموذج وتقدير معالمه ومعنوياته من خلال نتيجة جدول تحليل التباين لا بد من فحصه للتأكد من مدى تحقق افتراضات تحليل الانحدار باستخدام تحليل البواقي للكشف عن الافتراضات التالية:

- خطية دالة الانحدار (Linearity)
- ثبات تباين حدود الخطأ العشوائي (Homoscedasticity)
- استقلالية حدود الخطأ العشوائي(Independence)
- التوزيع الطبيعي لحدود الخطأ العشوائي(Normality) كما وردت في (إسماعيل،2001).

اختبار الافتراضات المتعلقة بحد الخطأ العشوائي:

أولا: الرسم البياني

ويمكن التحقق من توفر فرضيات النموذج الخطي البسيط من خلال تخطيط Scatter plots بتمثيل ($\hat{y}$)على المحور الأفقي أو(X) يقابله الخطأ العشوائي e أو الأخطاء المعيارية Standardized Residual يرمز لها(e_s)على المحور الرأسي أو (y). كما هو واضح في الشكل الآتي:



شكل (2)

الرسم البياني للمتغير المستقل مقابل الأخطاء العشوائية

يتبين من الشكل (2) أنه:

في (a) توفر فروض التحليل جميعها (عدم وجود مشكلة).

وفي (b) زيادة تباين الخطأ العشوائي بزيادة $\hat{y}$.

وفي (c) زيادة وتناقص في تباين الخطأ العشوائي (مشكلة عدم تجانس تباين الخطأ العشوائي).

وفي (d) عدم ملائمة العلاقة الخطية (يتوجب استعمال نماذج أخرى مثلا نموذج الدرجة الثانية) (بشير، 2003)

ثانيا: الاختبارات الإحصائية للافتراضات:

اختبار نقص المطابقة (lack of fit test)

إن أحد الافتراضات القياسية في تحليل الانحدار هو أن العلاقة بين المتغير التابع والمستقل في نموذج الانحدار الخطي البسيط علاقة خطية، ويتم اختبار هذا الافتراض من خلال رسم العلاقة بين المتغير التابع والمتغير المستقل عند عدم وجود تكرار لقيم المتغير المستقل، أما في حالة وجود تكرار لقيم المتغير المستقل فإن الرسم البياني لا يكشف عن العلاقة الخطية بين المتغير المستقل والتابع، وكذلك اختبارات المعنوية للنموذج الخطي، فلا بد من إجراء اختبار نقص المطابقة للتحقق من أن البيانات

مطابقة للنموذج الخطي بشكل جيد ويتم فيه تقسيم مجموع مربعات البواقي إلى:

1. مجموع مربعات الخطأ النقي(Pure error) ويرمز له بالرمز SS(p.e)

$$SS(p.e) = \sum_{j=1}^{r_i} (y_{ij} - \bar{y}_{i.})^2 \qquad (8)$$

بدرجة حرية $r_i - 1$

2. مجموع مربعات نقص المطابقة(Lack of Fit) ويرمز له بالرمز SS(L.O.F) وهو عبارة عن الفرق بين مجموع مربعات الخطأ ومجموع مربعات الخطأ النقي

$$SS(L.O.F) = SSe - SS(p.e) \qquad (9)$$

ولاختبار نقص المطابقة نستخدم الاختبار الإحصائي:

$$F = \frac{MS(L.o.F)}{MS(p.e)} \qquad (10)$$

ومن ثم يتم فحص معنوية (F) أو عدم معنويتها للتأكد من مدى مطابقة النموذج الخطي للبيانات من خلال اختبار الفرضية الصفرية والبديلة.

الفرضية الصفرية: النموذج الخطي ملائم للبيانات.

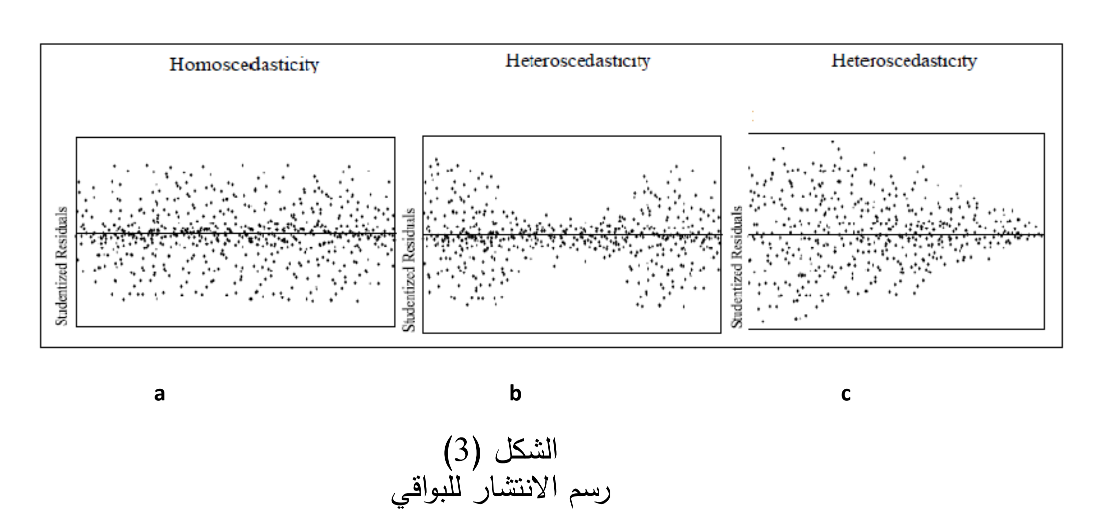
والفرضية البديلة: النموذج الخطي لا يلائم البيانات.

ويتم الحكم على معنوية أو عدم معنوية اختبار F من خلال تحليل جدول تحليل التباين للبواقي والذي يتم الحصول عليه من خلال البرنامج الإحصائي SPSS. فإذا تم رفض الفرضية الصفرية فيدل ذلك على أن النموذج الخطي غير ملائم للبيانات (بغض النظر عن نتيجة اختبار معنوية معالم معادلة الانحدار)، ويكون نموذج من الدرجة الثانية أو غيره يلائم البيانات(الراوي، 1987)

اختبار تجانس التباين للبواقي (Homoscedasticity test)

يعتبر افتراض تجانس التباين للبواقي من الافتراضات الضرورية والهامة لتقدير معالم نموذج الانحدار الخطي البسيط باستخدام طريقة المربعات الصغرى، وأن عدم تحقق هذا الافتراض ينتج عنه مشكلة تدعى عدم تجانس التباين للبواقي

(Heteroscedasticity) والتي تؤدي إلى أن يكون تباين معالم نموذج الانحدار المقدرة في حالة عدم تجانس التباين يختلف عن تباينها في حالة تحقق افتراض تجانس التباين ويترتب على هذا الاختلاف أن تكون النتائج التي نحصل عليها من الاستدلال الإحصائي (اختبارات المعنوية للمعالم وللنموذج وفترات الثقة والتنبؤ) غير دقيقة، ولا تمتاز بخاصية الكفاءة (Goldfeld and Quandt, 1965)، ويستخدم رسم شكل انتشار البواقي أو البواقي المعيارية مقابل المتغير المستقل للكشف عن تحقق افتراض تجانس التباين أو عدم تحققه والشكل (3) يبين رسم شكل الانتشار للبواقي.



a b c

الشكل (3)
رسم الانتشار للبواقي

يلاحظ من خلال الشكل (a–3) أن تباين البواقي تتوزع بشكل عشوائي ولا يعتمد توزيعها على المتغير المستقل مما يشير إلى تحقق افتراض تجانس البواقي، كما يلاحظ من الشكل (b–3) أن تباين البواقي يتناقص حتى يصل أقل حد له عند القيم المتوسطة للمتغير المستقل، ثم يبدأ في الزيادة تدريجيا حتى يصل إلى أعلى حد له عند القيم الكبيرة للمتغير المستقل مما يعني بأن تباين البواقي غير ثابت لجميع قيم المتغير المستقل، وهذا يشير إلى عدم تحقق افتراض تجانس التباين (Heteroscedasticity)، ويظهر الشكل (c–3) أن تباين البواقي يتناقص تدريجيا بزيادة قيم المتغير المستقل إلى أن يصل أقل حد له عند القيم الكبيرة للمتغير المستقل، وهذا يدل أيضا على عدم تحقق افتراض تجانس البواقي.

ويمكن الكشف عن مدى تحقق افتراض تجانس التباين للبواقي من خلال مجموعة من الاختبارات الإحصائية منها اختبار Goldfeld-Quandt، اختبار White، اختبار

Park Test اختبار Brattle's Test. ويعتبر اختبار Goldfeld-Quandt من الاختبارات الأكثر استخداما في الكشف عن مدى تحقق او انتهاك افتراض ثبات التباين ويتم من خلال الإجراءات الآتية:

1. ترتيب مشاهدات المتغير المستقل ترتيبا تصاعديا.
2. استبعاد 20% من المشاهدات الوسطى لكل من المتغير المستقل والتابع، ثم تكوين مجموعتين من المشاهدات بحيث يكون لكل مجموعة معادلة انحدار خاصة بها.

المجموعة الأولى: وتتمثل في المشاهدات الخاصة بكل من المتغير (X) والمتغير (Y) الواردة قبل المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي:

$$Y_{1i} = a + bX_{1i} + \varepsilon_{1i} \quad (11)$$

المجموعة الثانية: وتتمثل في المشاهدات الخاصة بكل من المتغير(X) والمتغير (Y) الواردة بعد المشاهدات التي تم استبعادها، والمعادلة الخاصة بهذه المجموعة هي:

$$Y_{2i} = c + dX_{2i} + \varepsilon_{2i} \quad (12)$$

- تم تقدير معاملات المعادلتين السابقتين باستعمال المربعات الصغرى باستخدام المعادلات:

$$\hat{Y}_{1i} = \hat{a} + \hat{b}X_{1i} \quad (13)$$
$$Y_{2i} = \hat{c} + \hat{d}X_{2i}$$

- الحصول على القيم المقدرة لحد الخطأ باستخدام المعادلات:

$$\hat{\varepsilon}_{1i} = Y_{1i} - \hat{Y}_{1i} \quad (14)$$
$$\hat{\varepsilon}_{2i} = Y_{2i} - \hat{Y}_{2i}$$

- إيجاد القيمة المحسوبة لإحصائية F باستخدام المعادلة:

$$F = \frac{\sum \hat{\varepsilon}_{2i}^2}{\sum \hat{\varepsilon}_{1i}^2} \quad (15)$$

- إيجاد درجات الحرية باستخدام المعادلة:

$$DF = \frac{n - m - 2(k)}{2} \quad (16)$$

حيث k : عدد المتغيرات المستقلة، m : عدد المشاهدات المستبعدة، n: عدد المشاهدات الكلية

- إيجاد القيمة المجدولة للإحصائي (F) عند درجات الحرية لكل من البسط والمقام، ومستوى معنوية معين.
- مقارنة القيم المحسوبة للإحصائي (F) والقيمة المجدولة له
- اختبار الفرضية الصفرية والبديلة.
- الفرضية الصفرية: لا توجد فروق ذات دلالة إحصائية في تباين البواقي عند مستوى دلالة محدد (تجانس التباين للبواقي).
- الفرضية البديلة: توجد فروق ذات دلالة إحصائية في تباين البواقي عند مستوى دلالة محدد (عدم تجانس التباين للبواقي)،

فإذا كانت قيمة (F) المحسوبة أكبر من قيمة (F) المجدولة نرفض الفرضية الصفرية، ونقبل الفرضية البديلة، ويكون تباين البواقي غير ثابت(عدم تجانس البواقي) (Heteroscedasticity)

أما إذا كانت قيمة (F) المحسوبة أقل من قيمة (F) المجدولة، يتم قبول الفرضية الصفرية ورفض الفرضية البديلة(عطيه، 2000)

اختبار استقلالية حدود الخطأ العشوائي(البواقي) (independence)

يعد افتراض استقلالية حدود الخطأ العشوائي (البواقي) من الافتراضات الأساسية لتقدير معالم نموذج الانحدار باستخدام طريقة المربعات الصغرى الاعتيادية، بمعنى أخر عدم وجود ارتباط ذاتي بين حدود الخطأ العشوائي، وعدم تحقق هذا الافتراض يؤثر سلباً على نتائج المربعات الصغرى ويؤدي إلى أن تكون النتائج التي نحصل عليها من الاستدلال الإحصائي مثل إحصاءات (F) ، (T)، (R^2) أكبر من قيمها الحقيقية، مما يجعل نموذج الانحدار الخطي الذي نحصل عليه غير دقيق، ويعود سبب عدم تحقق هذا الافتراض إلى:

- إهمال بعض المتغيرات المستقلة في النموذج المراد تقديره.
- الصياغة الرياضية الخاطئة للنموذج.
- عدم دقة البيانات المستخدمة في النموذج.

ويمكـن الكشـف عـن تحقـق افتـراض اسـتقلالية البـواقي مـن خـلال اختبـار (Durbin- Watson Test) الذي يعتبر من أهم الاختبارات الإحصائية التي تستخدم للكشف عن الارتباط الذاتي بين البواقي، ويهدف إلى اختبار الفرضيات التالية:

الفرضية الصفرية: لا يوجد ارتباط ذاتي بين البواقي عند مستوى الدلالة المحدد.

الفرضية البديلة: يوجد ارتباط ذاتي بين البواقي عند مستوى الدلالة المحدد.

$$H_0 : \rho = 0$$
$$H_1 : \rho \neq 0$$

ويتم اختبار فرضية العدم H_0 من خلال حساب إحصائية دربين واتسون DW، من خلال المعادلة:

$$DW = \frac{\sum_{t=2}^{n}\left(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}\right)^2}{\sum_{t=1}^{n}\hat{\varepsilon}_{t-1}^2} \qquad (17)$$

ويمكن كتابة الإحصائية أيضا بدلالة مقدر معامل الارتباط (ρ)

من خلال المعادلة:

$$DW \cong 2(1-\hat{\rho}) \qquad (18)$$

حيث إن الإحصائي DW يمثل القيمة المحسوبة للاختبار وتأخذ قيمها بين(0 و4) اعتماداً على القيمة المقدرة لمعامل الارتباط (بني هاني، 2014).

ويتم الحصول على قيمة DW من خلال برنامج SPSS ويتم قبول أو رفض الفرضية الصفرية من خلال مقارنة قيمة DW المحسوبة بقيمة DW الجدولية، وكقاعدة عملية إذا تراوحت قيمة DW الجدولية بين (1.5 إلى 2.5)فإننا نقبل الفرضية الصفرية التي تنص على عدم وجود ارتباط ذاتي بين البواقي(استقلالية البواقي) (نوروسيس، 2010).

اختبار التوزيع الطبيعي للبواقي (Normality):

من الافتراضات الأساسية لتحليل الانحدار الخطي أن تكون البواقي موزعة توزيعا طبيعيا، ويمكن أن يعود سبب عدم توزيع البواقي توزيعا طبيعيا إلى:

- توزيع المجتمع للمتغير التابع غير طبيعي.

- عدم ثبات التباين للبواقي.

ويؤدي انتهاك هذا الافتراض إلى أن تكون نتائج تحليل الانحدار غير دقيقة ولا يعتمد عليها في عملية التنبؤ، ويتم التحقق من هذا الافتراض من خلال مجموعة من الاختبارات الاحصائية منها اختبار مربع كاي ، اختبار Kolmogrov – smirnov ، اختبار Shapiro – Wilk ، ومن نتائج هذه الاختبارات التي سيتم الحصول عليها من برنامج SPSS يتم اختبار الفرضية الصفرية للتوزيع الطبيعي، حسب

الفرضية الصفرية: تنص على أن البواقي تتبع التوزيع طبيعي عند مستوى الدلالة المحدد.

الفرضية البديلة: تنص على أن البواقي لا تتبع التوزيع الطبيعي عند مستوى الدلالة المحدد.

القيم الشاذة (outlier values):

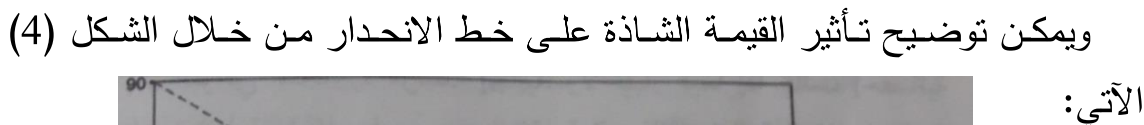
وهي القيم التي تبتعد كثيراً عن خط الانحدار ويكون حد الخطأ لها كبيرا مقارنة ببقية القيم الطبيعية الأخرى، ويكون لها تأثير كبير على النموذج الخطي ومعلماته (الراوي، 1987). وقد عرفها بارنت Barnett (1978) بأنها المشاهدة التي تبدو غير منطقية إذا قورنت بسائر مجموعة البيانات، وعرفها الجبوري AL-Jobouri (1976) بأنها تلك القيمة التي تكون غير منسجمة مع بقية بيانات المجموعة لمتغير من المتغيرات لظاهرة معينة أو مجموعة من الظواهر، أو أن القيم الشاذة هي القيم التي تأتي من مجتمع يختلف عن مجتمع الدراسة، وعرفها بروس Bross (1961) بأنها المشاهدة التي تظهر منحرفة بشكل كبير عن سائر مكونات العينة التي وجدت فيها تلك العينة، أما فريمان Freeman (1980) فقد عرفها بانها أي مشاهدة لم تتولد بالطريقة العامة التي ولدت الأغلبية العظمى من مشاهدات البيانات، كما عرف كيلر Keller (2000) المشاهدات الشاذة بأنها المشاهدات التي تقع بعيدة عن معادلة الانحدار، ويكون لها خطأ كبير مقارنة ببقية المشاهدات الطبيعية الأخرى في مجموعة البيانات ولذلك سيكون لها تأثير في النموذج الخطي وتقديراته.

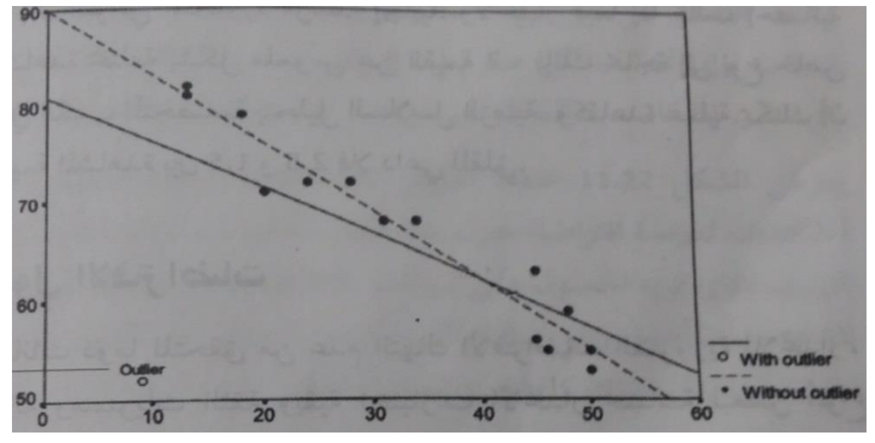
ويرجع وجود هذه القيم الشاذة إلى أخطاء إما في مرحلة تجميع البيانات أو في مرحلة إدخال البيانات في برامج التحليل الإحصائي على الحاسوب، وقد تكون هذه

القيم حقيقية ناتجة عن ظروف غير عادية.

وإن وجود قيم شاذة في بيانات نموذج الانحدار سواء في مشاهدات المتغير التابع أو المتغير المستقل يؤثر على تقديرات معالم النموذج والإحصاءات المرتبطة به خاصة إذا كان حجم العينة صغيرا، وهذا يؤدي إلى تغير كبير في قيم الميل والتقاطع لخط انحدار المربعات الصغرى عندما يتم معالجة القيم الشاذة، لهذا يجب فحص المشاهدات للتأكد من أنها لا تحتوي قيما شاذة قبل عملية إجراء تحليل الانحدار (إسماعيل، 2001).

ويمكن توضيح تأثير القيمة الشاذة على خط الانحدار من خلال الشكل (4) الآتي:





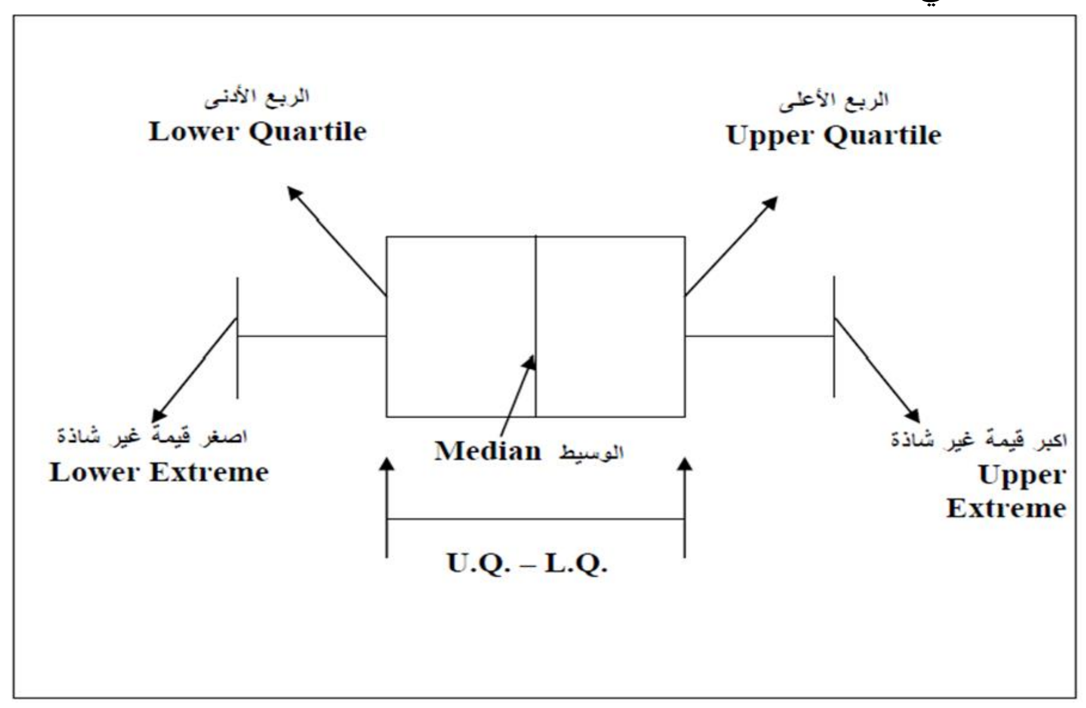
الشكل(4)

تأثير القيمة الشاذة على خط الانحدار

من خلال الشكل نلاحظ أن القيمة الشاذة الواقعة في الزاوية اليسرى السفلية غير منسجمة مع باقي القيم، وقد عملت على سحب خط الانحدار (الخط المتصل الممثل لجميع المشاهدات)، وبمعالجة القيمة الشاذة أدى ذلك إلى تغير في خط الانحدار (الخط المتقطع الممثل للمشاهدات عدا القيمة الشاذة)، وبالتالي حدث تغير كبير في معالم خط الانحدار (ميله ونقطة تقاطعه)، لهذا لا بد من الكشف عن القيم الشاذة ومعالجتها قبل عملية إجراء تحليل الانحدار (Nourusis, 1997)، وسوف تستخدم الدراسة مجموعة من الطرق للكشف عن القيم الشاذة ومنها الرسم الصندوقي (Boxplot)، وتحليل البواقي المعيارية.

طريقة الرسم الصندوقي (Boxplot):

اقترح توكي Tukey عام 1977 طريقة الصندوق والقطع المخططة مع الملخصات الخمسة The Box and Whisker Plots with 5 number summaries والموضحة بالشكل الآتي:



شكل (5)

مخطط الرسم الصندوقي

حيث يتكون الرسم الصندوقي من صندوق يمثل طوله المسافة بين قيمتي الربيع الأعلى (الثالث) (Q3) والربيع الأدنى (الأول) (Q1) وخط داخل الصندوق يمثل الوسيط Median وتقع 50% من المشاهدات داخل الصندوق، أي بين الربيع الأول والربيع الثالث، وتعتبر المشاهدات شاذة إذا كانت قيمها أكبر من Q3 + 1.5(Q3- Q1) ، أو أقل من Q1 -1.5(Q3-Q1) (السنجاري، 2001).

طريقة انتشار البواقي المعيارية:

يستخدم الرسم البياني لانتشار البواقي المعيارية مع القيم المتوقعة أو مع المتغير المستقل، فإذا أظهر شكل الانتشار بأن هناك مشاهدات تبعد بصورة واضحة عن بقية المشاهدات، فإن هذه المشاهدات تمثل مشاهدات شاذة لا بد من دراستها قبل إجراء تحليل الانحدار (نوروسيس، 2010).

معالجة القيم الشاذة

بعد التأكد من وجود القيم الشاذة بين المشاهدات، لا بد من معالجة هذه القيم وذلك من أجل الحصول على نموذج انحدار تكون نتائجه خالية من تأثير القيم الشاذة على معالمه المقدرة ، ويتم معالجة المشاهدات الشاذة بمجموعة من الطرق منها:

-حذف المشاهدات الشاذة إذا كان حجم العينة كبيرا وإعادة حل النموذج.

-استبدال القيم الشاذة بقيمة المتوسط المشذب Trimmed mean والتي تتلخص خطواته بترتيب المشاهدات التي تحتوي قيما شاذة تصاعديا أو تنازليا، وتحذف أكبر قيمة وأصغر قيمة، ثم إيجاد الوسط الحسابي للقيم المتبقية، أي ايجاد الوسط الحسابي المشذب والذي يمثل تقديرا للقيم الشاذة(حمودات، 2010).

2.2 الدراسات السابقة :

ستقوم هذه الدراسة بتحليل بيانات الانحدار الخطي البسيط للكشف عن انتهاكات افتراضات تحليل الانحدار من خلال الكشف عن القيم الشاذة وتأثيرها على افتراضات البواقي ودقة معالم معادلة الانحدار، الأمر الذي تطلب من الباحث الرجوع إلى معظم الدراسات ذات الصلة بالدراسة. ومن خلال البحث ومراجعة الدراسات السابقة ذات الصلة بموضوع الدراسة، تبين للباحث وجود مجموعة من الدراسات السابقة التي تناولت القيم الشاذة وطرق معالجتها ولكنها لم تتناول أثر هذه القيم الشاذة على الافتراضات الخاصة بالبواقي .

ومن هذه الدراسات:

قام ستيفن (Stevens, 1984) بدراسة هدفت إلى توضيح طرق الكشف عن القيم المتطرفة وتأثيرها على دقة معالم معاملات الانحدار، واستخدمت الدراسة عدة طرق للكشف عن القيم الشاذة المؤثرة في معادلة الانحدار ومنها مسافة كوك (cook) والبواقي ومسافة ماهلونبيز (Mahalanobis)

وتوصلت الدراسة إلى أن هناك قيم متطرفة تؤثر بشكل كبير على دقة معادلة الانحدار، وإن هناك قيم متطرفة لها تأثير بسيط على معادلة الانحدار.

كما قام اوسبورن ووترز (Osborne and Waters, 2002) بدراسة هدفت إلى

ضرورة اختبار الافتراضات في تحليل الانحدار الخطي المتعدد للوصول إلى معادلة انحدار متعدد تعبر بشكل دقيق عن العلاقة بين المتغير التابع والمتغيرات المستقلة، وتوصلت الدراسة إلى أن تحقق الافتراضات يساعد في تجنب الباحث من الوقوع في الخطأ من النوع الأول والخطأ من النوع الثاني، واوصت الدراسة بضرورة التحقق من افتراضات تحليل الانحدار قبل اعتماد معادلة الانحدار.

وقام اوسبورن واوفيربي (Osborne and Overbay , 2004) بدراسة هدفت إلى الكشف عن تأثير القيم المتطرفة على كل من اختبار (T) واختبار (F) وعلى معامل الارتباط، حيث بينت الدراسة تأثير القيم المتطرفة على كلا الاختبارين، وأوصت الدراسة بضرورة التعامل مع القيم المتطرفة ومعالجتها للتخلص من أثرها على الاختبارات.

وقام يونس وإبراهيم (2004) بدراسة هدفت إلى تحليل بيانات الانحدار الخطي التي تحتوي على تكرار في قيم المتغير المستقل لاكتشاف الخلل في فرض الخطية في نموذج الانحدار الخطي وعدم ملائمة البيانات للنموذج الخطي، وتوصلت الدراسة إلى أنه لابد من إجراء اختبار نقص المطابقة في مثل هذه الحالات، حيث أكدت نتيجة الاختبار أن هناك معادلة أفضل من المعادلة الخطية تلائم البيانات.

وقام دبدوب ويونس(2006) بدراسة هدفت إلى الكشف عن القيم الشاذة التي تشير إليها نتائج تحليل الانحدار والرسم الصندوقي بالإضافة إلى تأثير وجود ومعالجة القيم الشاذة على نتائج تحليل الانحدار. وتوصلت الدراسة إلى عدد من النتائج أهمها :أن القيم الشاذة تؤثر على تركيبة النموذج ونتائج تحليل الانحدار من خلال إدخال متغيرات غير مؤثرة على الظاهرة المدروسة واستبعاد بعض المتغيرات المؤثرة. جمعت بيانات هذه الدراسة من مستشفى البتول التعليمي للولادة في محافظة نينوى –العراق وخلال ثلاث فترات وأخذت عينة عشوائية بحجم 100 مولود لكل فترة.

أما دراسة قاسم و إسماعيل (2008) فهدفت إلى دراسة القيم الشاذة والكشف عنها بطريقة بيز (Bayesian) في نموذج تحليل الانحدار الخطي البسيط بأسلوب معاينة جبس والتي هي حالة خاصة لطريقة مونتي كارلو سلسلة ماركوف وتطبيق هذا الأسلوب على بيانات حقيقة تم الحصول عليها من أحد البحوث المنشورة لأجل الكشف عن القيم

الشاذة .وتوصلت الدراسة إلي عدد من النتائج أهمها :أن الكشف عن القيم الشاذة بأسلوب بيز باستخدام معاينة جبس يحل الكثير من المسائل المعقدة ولا سيما ذات النماذج المعقدة والصعبة؛ لأن معاينة جبس تقسم المسائل والنماذج المعقدة إلى مجموعة مسائل بسيطة يمكن تحليلها ومعالجتها بسهولة، كما أن النتائج تكون دقيقة، وذلك لان هذه الطريقة تعتمد على المعلومات الأولية للمعلمات والتي غالباً ما تكون متوفرة في معظم الدراسات، كما إن هذه الطريقة تمكننا من الحصول على احتمالية كون المشاهدة أو المشاهدات قيماً شاذة.

كما قام حمودات (2010) بدراسة هدفت إلى التعرف على تأثير وجود بعض القيم الشاذة على نتائج تحليل الانحدار مقارنة بمعالجتها ، حيث تم الكشف عن وجود القيم الشاذة بطريقة الرسم الصندوقي وعولجت بطريقة متوسط البتر. وتوصلت الدراسة إلى عدة نتائج أهمها أن وجود القيم الشاذة في بعض المتغيرات أثر سلباً على نتائج تحليل الانحدار حيث إنها قللت من قيمة (F) المحسوبة لمعادلة الانحدار التقديرية وكذلك قيمة معامل التحديد (R^2) وزادت من قيمة متوسط مربعات الخطأ مقارنةً مع نتائج التحليل بعد معالجة القيم الشاذة والتي أدت إلى عكس النتيجة. اعتمدت هذه الدراسات بيانات دراسة (مروان وفرح ، 2006) سابقة الذكر وكانت عينتها 100 مولود عن كل فترة تقويمية .

وقام يوسف (2015) بدراسة هدفت إلى التعرف على القيم الشاذة وتأثيرها في معالم معادلة الانحدار الخطي المتعدد، كما هدفت الدراسة إلى معالجة القيم الشاذة ومقارنة وجودها على معالم معادلة الانحدار الخطي المتعدد بطريقة الحذف وطريقة متوسط البتر وتوصلت الدراسة إلى أن طريقة حذف القيم الشاذة أدت إلى إعطاء نتائج أفضل في معادلة الانحدار الخطي المتعدد مقارنة بطريقة متوسط البتر، وتمثل مجتمع الدراسة في جميع الطلاب الذين تم قبولهم بكلية التربية في جامعة الجزيرة للعام 2009، وتم اختيار عينة الدراسة بحجم (30) طالبا وطالبة تم سحبهم عشوائيا بطريقة المعاينة العشوائية المنتظمة.

3.2 التعقيب على الدراسات السابقة

في ضوء ما تقدم من دراسات حول طرق الكشف عن القيم الشاذة وطرق معالجتها في تحليل الانحدار الخطي البسيط والمتعدد، تبين أن هذه الدراسات قد حاولت التوصل إلى أفضل نموذج يلائم معادلة الانحدار دون التحقق من افتراضات البواقي بالدرجة الأولى، واكتفت الدراسات بمعنوية النموذج ومعنوية الميل وقيمة مربع معامل التحديد والكشف عن القيم الشاذة للحكم على دقة معالم معادلة الانحدار، وجاءت هذه الدراسة للتحقق من افتراضات البواقي بعد ملائمة نموذج معادلة الانحدار الخطي البسيط.

الفصل الثالث

المنهجية والتصميم

يتضمن هذا الفصل وصفا لمجتمع الدراسة، وعينة الدراسة، وإجراءات الدراسة من جمع البيانات وتحليلها والتحقق من افتراضات الانحدار الخطي البسيط.

1.3 منهجية البحث:

تم إجراء بحث أساسي يهدف إلى استخدام البواقي والقيم الشاذة ومعالجتها للكشف عن انتهاكات افتراضات تحليل الانحدار الخطي البسيط، والتحقق من دقة معادلة الانحدار الخطي في بيانات عينة الدراسة، ويعتمد على المنهج الوصفي التجريبي لتحقيق أهداف الدراسة.

2.3 مجتمع الدراسة:

تكون مجتمع الدراسة من الأمثلة الواردة في المراجع المتخصصة في التحليل الإحصائي والدراسات التي اعتمدت تحليل الانحدار الخطي البسيط.

3.3 عينة الدراسة:

تم اختيار عينة عشوائية مؤلفة من بيانات (22) دراسة من مجتمع الدراسة، والتي استخدمت موضوع تحليل الانحدار الخطي البسيط.

4.3 إجراءات الدراسة:

تم في الفصل الرابع من فصول الدراسة التحليل الإحصائي لبيانات الدراسة للإجابة عن أسئلتها وفق الإجراءات الآتية:

1. إدخال بيانات الدراسات المستخدمة في الدراسة في البرنامج الإحصائي (SPSS)
2. عمل تحليل انحدار خطي بسيط لبيانات عينة الدراسة كما وردت في المراجع والدراسات التي أخذت منها البيانات، وعرض نتائج تحليل الانحدار (معامل

التحديد، اختبار F، معاملات نموذج الانحدار).

3. تم القيام بمجموعة من الاختبارات (نقص المطابقة، تجانس التباين، واستقلالية البواقي، والتوزيع الطبيعي) وعرض نتائجها.
4. الرسم البياني (Scatterplot) للبواقي المعيارية مقابل المتغير المستقل وعرض الأشكال.
5. تم الكشف عن القيم الشاذة في متغيرات عينة الدراسة باستخدام طريقة (Tukey) وعرض النتائج.
6. معالجة القيم الشاذة بطريقة التعويض بالمتوسط المشذب (Trimmed mean).
7. عمل تحليل انحدار خطي بسيط لبيانات عينة الدراسة بعد معالجة القيم الشاذة وعرض نتائج التحليل.
8. إجراء اختبارات (نقص المطابقة، تجانس التباين، واستقلالية البواقي، والتوزيع الطبيعي) بعد معالجة القيم الشاذة.
9. المقارنة بين نتائج التحليل ونتائج اختبارات افتراضات تحليل الانحدار قبل معالجة القيم الشاذة وبعده.

الفصل الرابع

عرض النتائج ومناقشتها والتوصيات

1.4 عرض النتائج ومناقشتها

يتناول هذا الفصل عرضا للنتائج التي توصلت إليها الدراسة، والتي هدفت إلى الكشف عن انتهاكات افتراضات تحليل الانحدار الخطي البسيط ودقة تقدير معالم معادلة الانحدار من خلال تحليل البواقي والقيم الشاذة لعينة الدراسة، وقد تم عرض نتائج الدراسة التي تم الحصول عليها من خلال البرنامج الإحصائي (SPSS) حسب أسئلة الدراسة باتباع منهجية منظمة تقوم على عرض السؤال، ثم عرض نتائج التحليل منظمة في جداول لتسهل على القارئ المقارنة بين النتائج، ثم التعليق على النتائج للتوصل إلى الإجابة على أسئلة الدراسة.

السؤال الأول: **هل تعتبر معنوية نموذج الانحدار الخطي البسيط ومعاملاته دليلا كافيا على تحقق افتراضاته؟**

للإجابة عن السؤال تم إجراء ما يلي:

أولاً: تم حساب معامل التحديد(R Square)، واختبار(F)، ومعاملات نموذج الانحدار لكل الدراسات (عينة الدراسة). والجدول (1) يبين نتائج ذلك.

جدول (1)

نتائج تحليل الانحدار(معامل التحديد، اختبار F ، معاملات نموذج الانحدار)

البيانات	R Square	F	Sig	β_0	Sig.	β_1	Sig.
1	0.926	596.535	0.000	4.210	0.000	2.532	0.000
2	0.941	418.391	0.000	34.423	0.006	8.674	0.000
3	0.996	3698.780	0.000	769.059	0.000	0.733	0.000
4	0.846	159.894	0.000	−9.389	0.000	1.703	0.000
5	0.639	61.858	0.000	0.371	0.000	0.066	0.000
6	0.856	136.477	0.000	701.773	0.221	0.711	0.000

البيانات	R Square	F	Sig	β_0	Sig.	β_1	Sig.
7	0.785	54.814	0.000	8.149	0.578	0.735	0.000
8	0.685	49.920	0.000	213.646	0.000	16.541	0.000
9	0.450	26.202	0.000	32.635	0.000	0.139	0.000
10	0.704	38.044	0.000	−11.615	0.357	0.666	0.000
11	0.426	35.601	0.000	17.544	0.072	0.701	0.000
12	0.901	127.847	0.000	−19.199	0.007	0.484	0.000
13	0.963	5113.111	0.000	−2.59	0.012	1.120	0.000
14	0.787	103.374	0.000	0.453	0.011	0.733	0.000
15	0.782	100.698	0.000	0.394	0.032	0.805	0.000
16	0.750	83.948	0.000	0.590	0.002	0.747	0.000
17	0.959	648.525	0.000	20.370	0.000	−0.798	0.000
18	0.389	6.355	0.030	13.971	0.000	−0.165	0.03
19	0.938	195.653	0.000	89.985	0.000	−0.697	0.000
20	0.874	69.603	0.000	23.783	0.005	0.755	0.000
21	0.971	949.572	0.000	−0.038	0.843	1.213	0.000
22	0.414	13.407	0.002	109.562	0.000	−1.125	0.002

يبين جدول(1) أن قيمة معامل التحديد لجميع الدراسات مرتفعة مما يعني أن المتغير المستقل يفسر قيمة مرتفعة من التباين في المتغير التابع ، كما أظهرت النتائج معنوية اختبار(F) لجميع الدراسات، ومعنوية معامل الانحدار(β_1) ومن خلال هذه النتائج وحدها يعتمد معظم الباحثين على معادلة الانحدار من أجل التنبؤ واتخاذ القرارات والتي يمكن أن تكون غير دقيقة إذا لم يتم التحقق من افتراضات تحليل الانحدار.

ثانيـاً: التحقـق مـن افتراضـات البـواقي مـن خـلال الاختبـارات: نقـص المطابقـة (lack of fit) تجانس التبـاين (Homoscedasticity)، الاستقلالية (Independence)، التوزيع الطبيعي (Normality) والجدول (2) يبين نتائج ذلك.

جدول (2)

اختبارات التحقق من افتراضات البواقي

Normality		Independence	Homoscedasticity		Lack of fit		البيانات
Sig.	Statistic	Durbin-Watson	F_{table}	F_{calc}	Sig.	F	
0.200	0.013	1.607	2.190	2.252	0.001	5.244	1
0.2000	0.078	1.766	3.180	9.112	0.220	1.574	2
0.2000	0.097	2.356	6.390	1.501	0.992	0.142	3
0.200	0.113	0.677	2.980	0.191	0.873	0.546	4
0.200	0.098	2.127	2.690	2.272	0.589	0.823	5
0.200	0.141	2.192	3.440	28.582	0.909	0.246	6
0.200	0.133	2.405	5.050	0.1089	0.558	1.149	7
0.057	0.171	2.255	3.440	0.2067	0.285	1.418	8
0.038	0.155	1.551	2.690	1.552	0.679	0.848	9
0.200	0.148	2.374	5.050	2.416	0.773	0.606	10
0.000	0.253	0.304	2.190	0.125	1.000	0.062	11
0.200	0.135	1.163	3.390	15.844	لا يوجد قيم مكررة في المتغير المستقل		12
0.006	0.077	2.016	1.530	0.669	0.000	3.376	13
0.200	0.109	1.337	2.980	0.123	0.885	0.377	14
0.200	0.095	1.373	2.980	0.553	0.866	0.460	15
0.200	0.102	1.608	2.980	2.970	0.618	0.940	16
0.200	0.125	1.711	2.980	0.768	0.539	0.879	17
0.200	0.102	0.794	9.280	0.314	0.006	11.508	18
0.200	0.147	1.983	6.390	1.381	0.575	0.933	19
0.200	0.152	2.450	9.280	0.140	0.531	1.753	20
0.200	0.126	2.489	2.980	10.736	0.560	1.629	21
0.200	0.131	2.143	4.280	2.176	0.089	2.533	22

يتبين من النتائج الواردة في جدول(2) ما يلي:

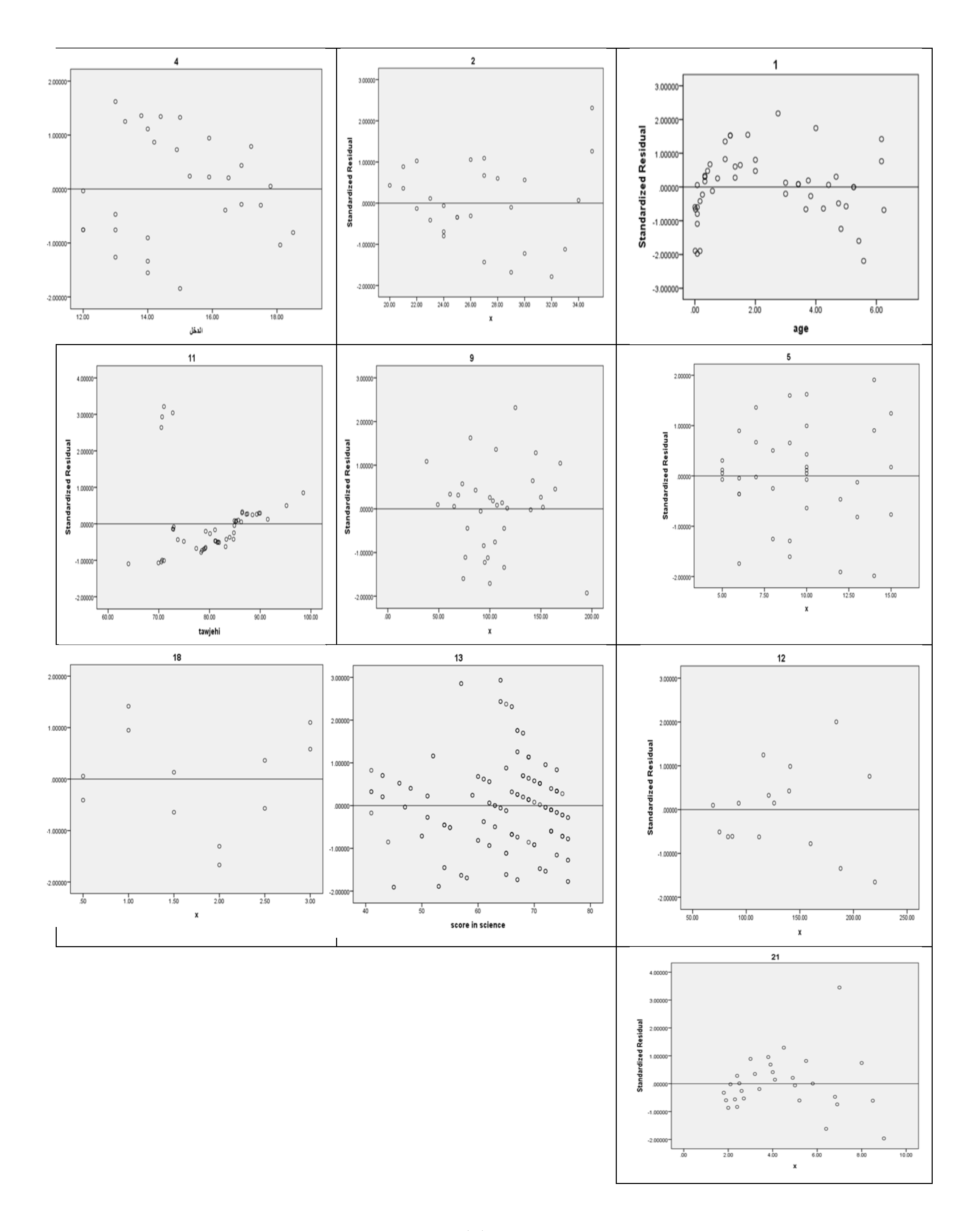
- بالنسبة لاختبار نقص المطابقة (Lack of fit) أظهر بأن قيمة P.Value للبيانات (1 ، 13 ، 18) أقل من مستوى المعنوية 0.05، وعليه فإن النموذج الخطي لا يلائم البيانات، وإن هناك معادلة من الدرجة الثانية او الثالثة تلائم البيانات بشكل اكثر دقة، وبالتالي فان الاعتماد على النموذج

الخطي لا يعبر عن العلاقة الدقيقة بين المتغير المستقل والمتغير التابع في كل منها.

- بالنسية لاختبار تجانس التباين (Homoscedasticity) أظهر أن البيانات (1، 2 ، 6، 12 ، 21) لا تحقق شرط تجانس التباين للبواقي حيث كانت قيم (F) المحسوبة أكبر من قيم (F) المجدولة.
- بالنسبة لاختبار الاستقلالية (Independence) أظهر أن قيمة اختبار (Durbin-Watson) للبيانات (4 ، 11 ، 18) أقل من (1.5) وعليه فإن البواقي لا تحقق شرط الاستقلالية وإنما يوجد ارتباط ذاتي بينها.
- بالنسبة لاختبار التوزيع الطبيعي (Normality) أظهر أن قيمة P.Value للبيانات (9 ، 11 ، 13) أقل من مستوى المعنوية 0.05، وعليه فإن البواقي لم تحقق شرط التوزيع الطبيعي.

من خلال النتائج تبين أن البيانات (1 ، 2 ، 4 ، 6 ، 9 ، 11 ، 12 ، 13 ، 18 ، 21) لم تحقق جميع شروط تحليل الانحدار الخطي وبالتالي فإن معادلة النموذج الخطي غير دقيقة ولا تعبر عن العلاقة الصحيحة بين المتغير التابع والمتغير المستقل، وأن القرارات المبنية عليها تكون مضللة.

ثالثاً: الرسم البياني للبواقي المعيارية مقابل المتغير المستقل والشكل (6) يبين نتائج ذلك.



شكل (6)

الرسم البياني للبواقي المعيارية مقابل المتغير المستقل

تبين من الشكل (6) بأن البيانات لم تحقق افتراضات الانحدار الخطي وهذا يدعم نتائج الاختبارات الواردة في جدول (2).

السؤال الثاني: **ما أثر القيم الشاذة ومعالجتها على دقة معادلة الانحدار الخطي البسيط ؟**

للإجابة عن السؤال تم إجراء ما يلي:

اولاً: الكشف عن القيم الشاذة في المتغير المستقل والمتغير التابع وفي البواقي المعيارية باستخدام طريقة Tukey من خلال البرنامج الإحصائي (SPSS) والجدول (3) يبين ذلك.

جدول (3)

القيم الشاذة

القيم الشاذة باستخدام طريقة Tukey

البيانات	المتغير المستقل	المتغير التابع	البواقي المعيارية
3			9
6			22،25،21
7			11
9			26
10			16
11			45،46،47،48
13	1،22،103،110، 131	22،45،85،97،103،154،1 94،131	15،61،84،104،124،170،1 93
14			21
15			9
16	21		15،26
20	9	9	-
21			29
22	18	18	19

أظهرت نتائج الكشف عن القيم الشاذة في عينة الدراسة بوجود قيم شاذة في كل من المتغير المستقل والمتغير التابع والبواقي المعيارية والمؤثرة على نتائج تحليل الانحدار الخطي.

ثانياً: تم إعادة التحليل للبيانات التي احتوت على قيم شاذة بعد معالجتها بطريقة التعويض بالمتوسط المشذب (Trimmed mean) والجدول(4) يبين معامل التحديد(R Square)، واختبار (F) ، ومعاملات نموذج الانحدار بعد معالجة القيم الشاذة.

جدول (4)

نتائج تحليل الانحدار(معامل التحديد، اختبار F ، معاملات نموذج الانحدار) بعد معالجة القيم الشاذة

البيانات	R Square	F	Sig	β_0	Sig.	β_1	Sig.
3	0.998	5514.398	0.000	734.827	0.000	0.774	0.000
6	0.936	334.576	0.000	1236.645	0.011	0.747	0.000
7	0.875	106.273	0.000	18.790	0.061	0.607	0.000
9	0.505	32.647	0.000	32.944	0.000	0.132	0.000
10	0.782	57.440	0.000	−14.408	0.184	0.693	0.000
11	0.959	1111.817	0.000	−12.553	0.000	1.052	0.000
13	0.975	7435.825	0.000	−3.267	0.000	1.127	0.000
14	0.880	204.682	0.000	0.484	0.000	0.770	0.000
15	0.860	171.545	0.000	0.204	0.178	0.881	0.000
16	0.857	167.642	0.000	0.567	0.000	0.761	0.000
20	0.967	823.854	0.000	20.414	0.000	−0.813	0.000
21	0.984	1636.559	0.000	0.119	0.412	1.160	0.000
22	0.614	30.234	0.000	111.396	0.000	−1.310	0.000

يلاحظ من خلال نتائج الجدول (4) ومقارنتها مع نتائج التحليل في الجدول (1) ارتفاع قيمة معامل التحديد (R Square) لجميع الدراسات بعد معالجة القيم الشاذة ،كذلك ارتفاع قيم اختبار (F) لجميع الدراسات والتغير في معاملات نموذج الانحدار وأن السبب في ارتفاع قيم اختبار (F) يعود إلى معالجة القيم الشاذة حيث إن قيمة (F) ناتجة من قسمة متوسط مجموع مربعات الانحدار على متوسط مجموع مربعات البواقي

ونتيجة لمعالجة القيم الشاذة أدى ذلك إلى تقليل مجموع مربعات البواقي وارتفاع في مجموع مربعات الانحدار مما سبب ارتفاع في قيم اختبار (F).

ثالثا: تم حساب مجموع مربعات البواقي قبل وبعد معالجة القيم الشاذة والجدول (5) يبين نتائج ذلك.

جدول (5)

مربعات البواقي قبل وبعد معالجة القيم الشاذة

البيانات	مجموع مربعات البواقي قبل معالجة القيم الشاذة	مجموع مربعات البواقي بعد معالجة القيم الشاذة
3	61482.190	41286.759
6	1906149.542	850131.913
7	36.122	20.769
9	29.454	26.560
10	15.098	9.073
11	35.095	2.529
13	18.985	2.772
14	0.065	0.037
15	0.066	0.043
16	0.077	0.044
20	0.180	0.143
21	0.188	0.110
22	107.905	71.025

يتبين من النتائج الواردة في الجدول (5) انخفاضا في مجموع مربعات البواقي بعد معالجة القيم الشاذة لذلك فإن للقيم الشاذة تأثيرا كبيرا على دقة معادلة الانحدار الخطي، وبالتالي فإن معادلة النموذج الخطي الناتجة بعد معالجة القيم الشاذة أكثر كفاءة ودقة من معادلة النموذج الخطي قبل معالجة القيم الشاذة.

السؤال الثالث: **ما أثر القيم الشاذة ومعالجتها على افتراضات بواقي تحليل الانحدار؟**

للإجابة عن السؤال تم إجراء الاختبارات التالية:

نقص المطابقة (Lack of fit)، تجانس التباين (Homoscedasticity)، الاستقلالية (Independence)، التوزيع الطبيعي (Normality) بعد معالجة القيم الشاذة والجدول (6) يبين نتائج ذلك.

جدول (6)

اختبارات التحقق من افتراضات البواقي

Normality		Independence	Homoscedasticity		Lack of fit		البيانات
Sig.	Statistic	Durbin–Watson	F_{table}	F_{calc}	Sig.	F	
0.200	0.089	2.500	6.390	1.501	0.945	0.286	3
0.200	0.129	2.325	3.440	5.741	0.233	0.922	6
0.200	0.101	2.369	5.050	0.221	0.110	50.431	7
0.306	0.157	1.513	2.690	1.244	0.718	.0757	9
0.200	0.120	2.417	5.050	1.157	0.570	0.939	10
0.151	0.113	2.125	2.190	0.103	0.855	0.256	11
0.064	0.061	1.453	1.530	0.661	0.000	3.758	13
0.200	0.078	1.436	2.980	0.243	0.968	0.197	14
0.200	0.093	1.883	2.980	1.282	0.962	0.269	15
0.200	0.109	1.355	2.980	1.911	0.811	0.562	16
0.200	0.105	1.810	2.980	0.421	0.801	0.556	20
0.200	0.055	1.632	2.980	3.494	0.707	0.871	21
0.132	0.166	1.717	4.280	0.560	0.280	1.492	22

أظهرت نتائج الاختبارات الأربعة الواردة في جدول(6)

- بالنسبة لاختبار نقص المطابقة (Lack of fit) أظهر بأن قيمة P.Value لبيانات الدراسة (13) أقل من مستوى المعنوية 0.05، وبناء عليه فإن النموذج الخطي لا يلائم هذه البيانات بعد معالجة القيم الشاذة وذلك لان العلاقة بين المتغير المستقل والمتغير التابع علاقة غير خطية.
- بالنسبة لاختبار تجانس التباين (Homoscedasticity) أظهر أن قيمة (F) المحسوبة قلت في البيانات (6 ، 21)، وهذا يدل على ان وجود القيم الشاذة في البيانات يساهم في انتهاك افتراض تجانس التباين للبواقي.

- بالنسبة لاختبار الاستقلالية (Independence) أظهر أن قيمة اختبار Durbin–Watson في البيانات (11) قبل معالجة القيم الشاذة كانت يساوي (0,304) والذي يشير إلى عدم استقلالية البواقي، وبعد معالجة القيم الشاذة في البيانات (11) أصبحت القيمة تساوي (2.125) والتي تشير إلى استقلالية البواقي وهذا يدل على أن معالجة القيم الشاذة ساهمت في تحقيق افتراض استقلالية البواقي.
- بالنسبة لاختبار التوزيع الطبيعي (Normality) أظهر أن قيمة P.Value للبيانات (9 ، 11 ، 13) أقل من مستوى المعنوية 0.05، قبل معالجة القيم الشاذة، وبعد معالجة القيم الشاذة للبيانات (9 ، 11 ، 13)، أصبحت قيمة P.Value أكبر من مستوى المعنوية 0.05، وعليه فإن البيانات حققت شرط التوزيع الطبيعي.

من خلال النتائج السابقة يتبين أن معالجة القيم الشاذة أدت إلى تحقيق افتراض استقلالية البواقي وافتراض التوزيع الطبيعي للبواقي، كما أظهرت نتائج الدراسة أن وجود القيم الشاذة كان سبباً من الأسباب التي أدت إلى عدم تحقق افتراض ثبات التباين وافتراض نقص المطابقة.

2.4 التوصيات

من خلال النتائج التي توصلت اليها الدراسة فهي توصي بما يلي:

1. التحقق من جميع افتراضات تحليل الانحدار الخطي البسيط وعدم الاعتماد على نتائج تحليل الانحدار فقط في الدراسات.
2. إجراء اختبار نقص المطابقة للبيانات في حالة وجود تكرار في قيم المتغير المستقل للتحقق من وجود العلاقة الخطية، وعدم الاعتماد على شكل الانتشار بين المتغيرين.
3. معالجة القيم الشاذة قبل إجراء تحليل الانحدار.
4. إجراء دراسة مقارنة بين طرق معالجة القيم الشاذة لتحديد أفضل طرق معالجة للقيم الشاذة للمساهمة في تحقيق افتراضات تحليل الانحدار الخطي.

قائمة المراجع

(أ) المراجع العربية

إسماعيل، محمد (2001). **تحليل الانحدار الخطي البسيط**، معهد الإدارة العامة، الرياض

بشير، سعد (2003). **دليلك إلى البرنامج الإحصائي SPSS** . الجهاز المركزي للإحصاء، جمهورية العراق.

بني هاني، عبدالرزاق (2014). **الاقتصاد القياسي نظرية الانحدار البسيط والمتعدد.** دار وائل للنشر والتوزيع: الاردن

الجبوري، شلال(1990).أهمية طريقة اكتشاف وتقدير القيم الشاذة في حالة الانحدار الخطي البسيط. **مجلة كلية الادارة والاقتصاد**. العدد الثاني، ص 316

حمودات، ألاء (2010) . تأثير القيم الشاذة ونقطة الأصل علي نتائج تحليل الانحدار المتعدد، **مجلة تكريت للعلوم الصرفة**، المجلد 15 ، العدد 1.

دبدوب، مروان ويونس، فرح. (2006)، تأثير القيم الشاذة على نتائج تحليل الانحدار مع التطبيق على المواليد الخدج، **مجلة علوم الرافدين**، المجلد 17العدد 1، ص،ص:62-81

الراوي، خاشع (1987). **المدخل إلى تحليل الانحدار**. مؤسسة دار الكتب للطباعة والنشر. جامعة الموصل. العراق

الزعبي، محمد والطلافحه، عباس (2012). **النظام الاحصائي SPSS**. ط3. دار وائل للنشر والتوزيع. الاردن

السنجاري، محمد (2001). **معالجة القيم المفقودة وتأثرها بالقيم الشاذة في تصميم العبور البسيط**. رسالة ماجستير. كلية علوم الحاسبات والرياضيات. جامعة الموصل، العراق.

السيفو، وليد، شلوف، فيصل وجواد، صائب (2006). **مشاكل الإقتصاد القياسي التحليلي**. الاهلية للنشر والتوزيع: عمان، الاردن.

السيفو، وليد، مشعل، احمد (2003). **الاقتصاد القياسي التحليلي بين النظرية والتطبيق**. دار مجدلاوي للنشر والتوزيع: عمان، الاردن

العتوم، شفيق (2015). **طرق الاحصاء باستخدام SPSS**. دار المناهج للنشر والتوزيع: عمان، الاردن.

عطيه؛ عبدالقادر (2000). **الاقتصاد القياسي بين النظرية والتطبيق**. ط2. الدار الجامعية للنشر والتوزيع والطباعة: الاسكندرية.

قاسم، محمد وإسماعيل، يونس (2008)، الكشف عن القيم الشاذة بأسلوب بيز باستخدام معاينة جبس، **المجلة العراقية للعلوم الإحصائية**، العدد (14).

نوروسيس، ماريجا (2010). **تحليل البيانات باستخدام 17 SPSS** . لجنة الترجمة والتأليف. شعاع للنشر والعلوم. سورية

يوسف، عصام (2015). **تأثير القيم الشاذة في معلمات نموذج الانحدار الخطي المتعدد**. جامعة السودان للعلوم والتكنولوجيا

يونس، عادل وابراهيم، موسى (2004). **اختبار نقص المطابقة في البيانات المكررة.** قسم الإحصاء التطبيقي- كلية العلوم- جامعة السودان للعلوم والتكنولوجيا.

(ب) المراجع الاجنبية

AL-Jobouri, S. (1976), **Test of Outliers Unpublished, M.s.c Thesis**. University of Baghdad.

Barnett, V. & Lewis, T. (1978), **Outliers in Statistical Data**, Jon Wiley & sons, New York.

Bross, J (1961), Outliers in Patternend Experiments: strategic Re-Appraisal. Technimetrics. 3,19-102.

Freeman, R (1980), **On the Number of Outliers in Data from a linear Model In Bayesian Statistics**, Ed .J.M.

Goldfeld, M. & Quandt. E, (1965), Some Test for Homoscedasticity. **Journal of the American Statistical Association**, 60pp.539-547.

Keller, G. & Brian Warrack, (2000), **Statistic For Management and Economics**, 5th edition,

Nourusis, M. (1997). **SPSS guide to data analysis**. New jersey, prentice – Hall.

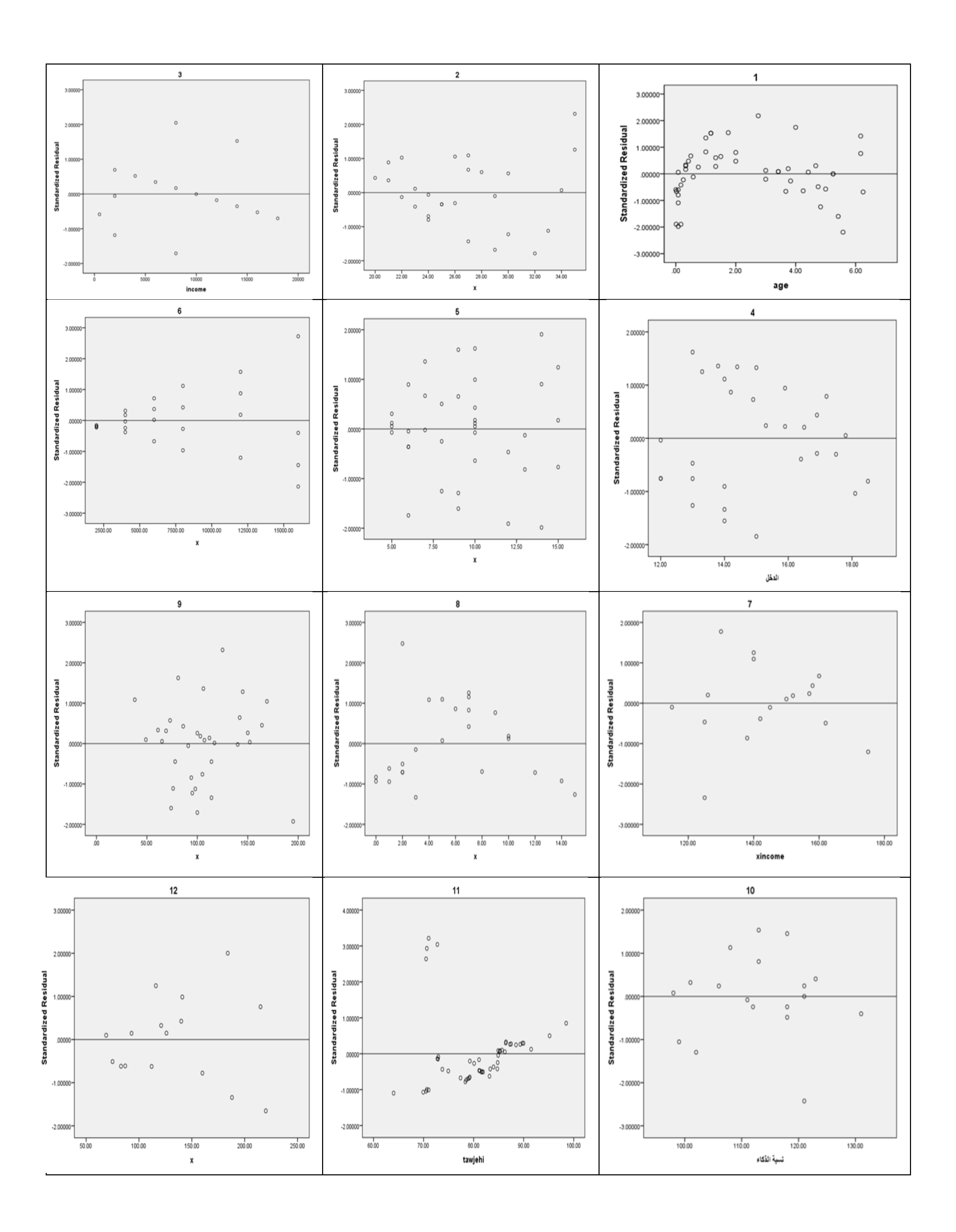
Osborne & Waters (2002). Four assumptions of multiple regression that researchers should always test. **Practical Assessment, Research & Evaluation**, 8(2)

Osborne, Jason. & Overbay, Amy (2004). The power of outliers and why researchers should always check for them. **Practical Assessment, Research & Evaluation**, 9(6)

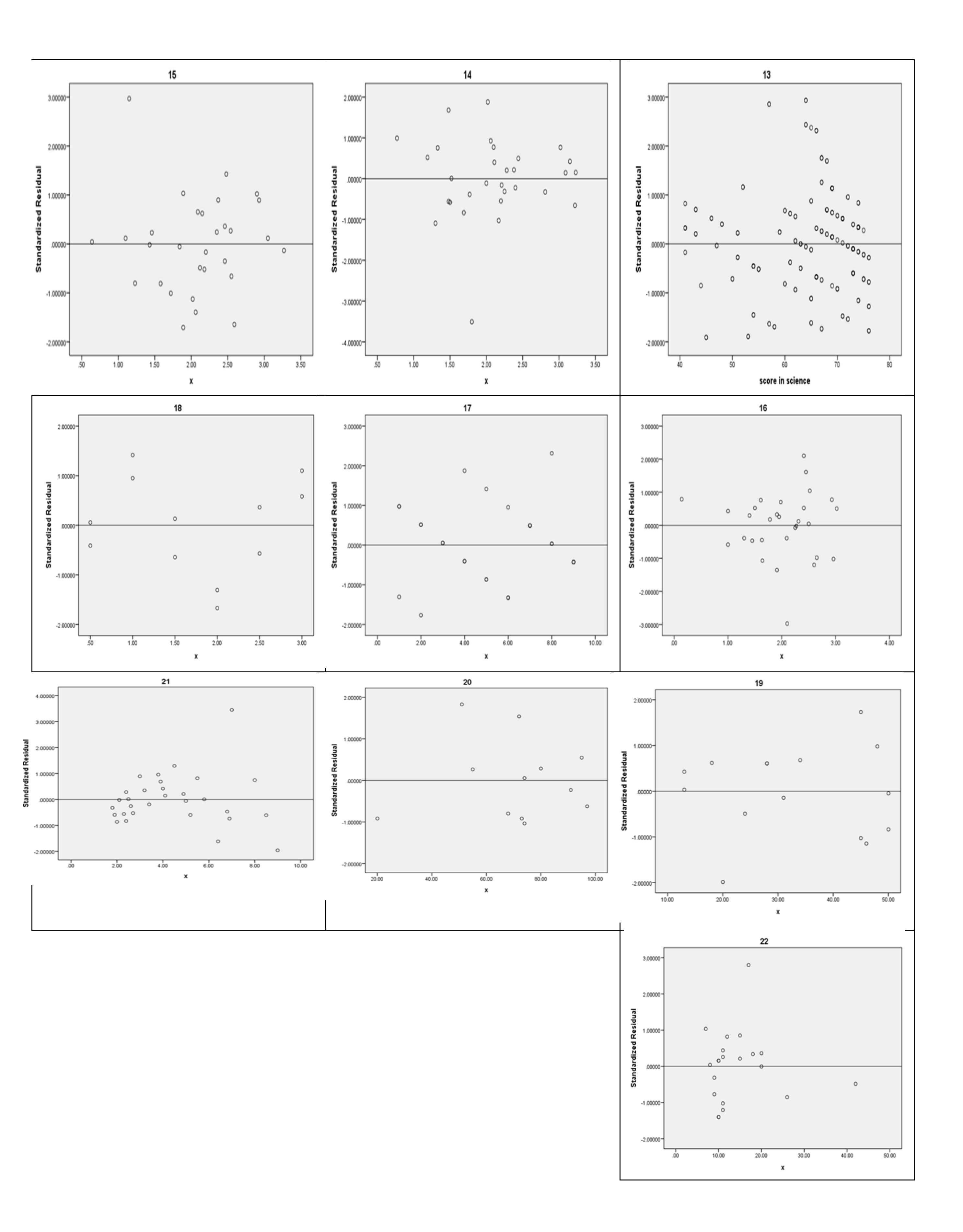
Stevens, J. P. (1984). Outliers and influential data points in regression analysis. **Psychological Bulletin**, *95*, 334-344.

Tukey, W.(1977), **Exploratory Data Analysis**, Addison-Wesley, reading, MA. Wiley and Sons, New York.

الملحق (أ)
رسم الانتشار للبواقي لجميع الدراسات



3
2
1
6
5
4
9
8
7
12
11
10
Standardized Residual
income
x
age
xincome
tawjehi
الدخل
نسبة الذكاء


15
14
13
18
17
16
21
20
19
22
Standardized Residual
x
score in science

الملحق (ب)

جدول توزيع F

F Values for = 0:05

					d_1				
d_2	1	2	3	4	5	6	7	8	9
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5
2	18.51	19.00	19.16	19.25	19.3	19.33	19.35	19.37	19.38
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96
inf	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88

F Values for = 0:05

d_1

d_2	10	12	15	20	24	30	40	60	120	inf
1	241.9	243.9	245.9	248.0	249.1	250.1	251.1	252.2	253.3	254.3
2	19.4	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.5
3	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.36
6	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	1.91	1.83	1.75	1.66	1.10	1.55	1.50	1.43	1.35	1.25
inf	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

الملحق (ج)

بيانات الدراسات

1		2		3		4		5	
x	**y**	**x**	**y**	**x**	**y**	**x**	**y**	**x**	**y**
3.00	11.50	20.00	212.00	500	1000	15.00	18.00	5.00	.69
5.00	16.00	21.00	220.00	2000	2300	16.90	19.00	15.00	1.39
.50	6.50	21.00	225.00	2000	2500	16.50	19.00	12.00	1.09
4.00	17.00	22.00	235.00	2000	2000	15.90	19.00	9.00	.76
1.33	8.50	22.00	224.00	4000	4000	18.10	20.00	12.00	.86
1.00	8.80	23.00	230.00	6000	5500	17.50	20.00	10.00	1.04
6.17	22.00	23.00	235.00	8000	6500	16.90	20.00	13.00	1.10
3.42	13.00	24.00	235.00	8000	7000	18.50	21.00	10.00	1.02
3.67	12.50	24.00	242.00	8400	8400	17.80	21.00	10.00	1.10
5.42	15.50	24.00	236.00	10000	8500	17.20	21.00	9.00	1.07
1.17	9.50	25.00	248.00	12000	10000	13.30	15.00	7.00	.94
4.42	15.50	25.00	248.00	14000	11500	13.00	15.00	14.00	.98
1.17	9.50	26.00	257.00	14000	12000	14.20	16.00	5.00	.71
2.75	14.50	26.00	270.00	16000	13000	14.00	16.00	5.00	.75
6.25	19.00	27.00	255.00	18000	14500	13.80	16.00	10.00	1.05
1.50	9.00	27.00	279.00			15.30	17.00	14.00	1.44
4.25	14.00	27.00	275.00			14.90	17.00	6.00	.76
2.00	10.50	28.00	283.00			14.40	17.00	5.00	.72
.42	6.00	29.00	270.00			16.40	18.00	10.00	1.29
5.58	15.00	29.00	285.00			15.90	18.00	6.00	.71
3.42	13.00	30.00	283.00			12.00	11.00	8.00	.86
6.17	21.00	30.00	300.00			12.00	10.00	10.00	.93
3.00	12.00	32.00	295.00			12.00	10.00	6.00	.91
5.25	17.50	33.00	310.00			13.00	12.10	8.00	.70
.33	5.50	34.00	330.00			13.00	11.70	15.00	1.56
.33	5.30	35.00	360.00			13.00	11.00	10.00	1.06
.75	6.50	35.00	350.00			14.00	13.20	9.00	.71
3.83	13.50					14.00	12.60	10.00	1.19
.25	4.50					14.00	12.30	7.00	.83
4.75	15.50					15.00	13.60	6.00	.49
4.67	16.50							15.00	1.24
1.75	11.00							9.00	1.22
5.25	17.50							8.00	.98
4.83	14.55							6.00	.71
2.00	10.00							14.00	1.60
.17	4.00							13.00	1.21
.08	3.50							7.00	1.05
1.00	8.00								
1.33	8.00								
3.75	14.00								
.17	1.75								
.08	3.20								
.33	5.55								
.08	2.75								
.01	1.35								
.58	5.50								
.08	4.50								
.02	3.25								
.00	3.30								
.08	1.40								

6		7		8		9		10	
x	**y**	**x**	**y**	**x**	**y**	**x**	**y**	**x**	**y**
2000.00	1850.00	158.00	127.00	2.00	212.00	38.00	44.00	98.00	54.00
2000.00	1800.00	160.00	130.00	5.00	300.00	49.00	40.00	99.00	50.00
2000.00	1850.00	175.00	129.00	15.00	400.00	61.00	43.00	101.00	57.00
2000.00	1900.00	157.00	125.00	1.00	200.00	65.00	42.00	102.00	51.00
4000.00	3200.00	130.00	115.00	3.00	256.00	69.00	44.00	106.00	60.00
4000.00	3500.00	150.00	119.00	.00	173.00	73.00	46.00	108.00	65.00
4000.00	3800.00	152.00	121.00	7.00	391.00	74.00	34.00	111.00	62.00
4000.00	3000.00	162.00	124.00	12.00	377.00	76.00	37.00	112.00	62.00
4000.00	4000.00	125.00	97.00	2.00	212.00	78.00	41.00	113.00	67.00
6000.00	4000.00	115.00	92.00	6.00	355.00	81.00	53.00	113.00	70.00
6000.00	5000.00	125.00	85.00	10.00	385.00	86.00	47.00	118.00	66.00
6000.00	5500.00	145.00	114.00	8.00	312.00	91.00	45.00	118.00	73.00
6000.00	6000.00	140.00	118.00	2.00	368.00	94.00	41.00	118.00	65.00
8000.00	5000.00	142.00	110.00	2.00	222.00	95.00	39.00	121.00	69.00
8000.00	6000.00	140.00	119.00	7.00	350.00	98.00	40.00	121.00	70.00
8000.00	7000.00	138.00	104.00	.00	168.00	100.00	37.00	121.00	59.00
8000.00	8000.00	126.00	102.00	14.00	400.00	100.00	48.00	123.00	72.00
12000.00	7500.00			7.00	386.00	103.00	48.00	131.00	74.00
12000.00	9500.00			9.00	400.00	105.00	43.00		
12000.00	10500.00			1.00	184.00	106.00	55.00		
12000.00	11500.00			5.00	350.00	107.00	48.00		
16000.00	9000.00			7.00	370.00	112.00	49.00		
16000.00	11500.00			4.00	333.00	114.00	46.00		
16000.00	10000.00			10.00	388.00	114.00	41.00		
16000.00	16000.00			3.00	198.00	117.00	49.00		
						125.00	63.00		
						140.00	52.00		
						142.00	56.00		
						145.00	60.00		
						150.00	55.00		
						152.00	54.00		
						164.00	58.00		
						169.00	62.00		
						195.00	49.00		

11		12		14		15		16	
x	**y**	**x**	**y**	**x**	**y**	**x**	**y**	**x**	**y**
55.82	64.01	75.00	13.00	3.09	2.88	3.05	2.88	2.45	2.88
60.12	69.93	69.00	15.00	2.20	2.01	2.18	2.01	1.61	2.01
60.69	70.50	83.00	16.00	2.28	2.27	2.55	2.27	1.98	2.27
61.08	70.65	87.00	18.00	3.22	2.77	2.48	2.77	2.52	2.77
61.26	71.01	93.00	27.00	2.21	2.12	2.20	2.12	1.95	2.12
67.75	72.75	112.00	30.00	2.10	2.28	2.46	2.28	2.27	2.28
67.77	72.88	121.00	42.00	1.48	2.04	2.59	2.04	2.09	2.04
68.21	72.95	116.00	47.00	2.81	2.54	2.37	2.54	2.41	2.54
66.65	73.75	126.00	43.00	2.25	2.11	1.15	2.11	1.91	2.11
67.20	74.94	140.00	52.00	2.40	2.25	2.09	2.25	2.25	2.25
67.75	77.39	141.00	57.00	2.11	2.19	1.89	2.19	2.60	2.19
67.77	78.35	160.00	52.00	3.15	3.00	2.90	3.00	2.93	3.00
68.21	78.52	188.00	61.00	2.44	2.47	2.46	2.47	2.50	2.47
68.67	78.88	220.00	74.00	2.02	2.51	2.54	2.51	2.96	2.51
68.94	79.12	215.00	91.00	3.23	2.99	3.27	2.99	2.41	2.99
69.19	79.23	184.00	86.00	2.38	2.35	2.35	2.35	2.31	2.35
71.90	79.25			3.02	2.99	2.93	2.99	3.02	2.99
72.10	80.10			2.06	2.29	2.15	2.29	2.65	2.29
73.42	81.08			1.30	1.17	1.23	1.17	1.00	1.17
71.60	81.11			1.50	1.46	1.89	1.46	1.00	1.46
71.67	81.24			1.80	.92	.64	.92	.14	.92
71.67	81.55			2.17	1.86	1.84	1.86	1.50	1.86
71.85	81.70			1.69	1.54	1.43	1.54	1.45	1.54
71.90	81.89			1.77	1.72	2.02	1.72	1.40	1.72
72.10	83.17			1.19	1.51	1.72	1.51	1.64	1.51
73.42	83.32			.77	1.31	1.10	1.31	2.10	1.31
74.23	83.98			2.00	1.97	2.12	1.97	1.78	1.97
74.43	84.74			1.48	1.45	1.58	1.45	1.30	1.45
75.54	84.81			1.33	1.68	2.06	1.68	1.63	1.68
76.82	84.90			1.52	1.63	1.46	1.63	1.91	1.63
77.62	84.91								
77.67	85.17								
77.87	85.29								
78.24	85.68								
78.36	86.25								
79.96	86.42								
80.18	86.46								
80.34	87.26								
80.56	87.49								
81.09	88.47								
81.79	89.32								
82.26	89.75								
82.42	89.97								
82.47	91.48								
82.97	70.50								
84.83	70.65								
86.79	71.01								
86.97	72.75								
87.33	95.18								
91.77	98.49								

17		18		19		20		21	
x	**y**	**x**	**y**	**x**	**y**	**x**	**y**	**x**	**y**
2.00	19.00	.50	13.90	31.00	68.00	51.00	74.00	2.00	2.00
5.00	16.00	.50	13.81	50.00	53.00	68.00	70.00	3.90	5.00
4.00	17.00	1.00	14.08	18.00	79.00	72.00	88.00	4.00	5.00
6.00	15.00	1.00	13.99	28.00	72.00	97.00	93.00	1.90	2.00
8.00	14.00	1.50	13.75	13.00	82.00	55.00	67.00	1.80	2.00
4.00	18.00	1.50	13.60	34.00	68.00	73.00	73.00	2.30	2.50
5.00	16.00	2.00	13.32	45.00	63.00	95.00	99.00	2.40	2.50
6.00	16.00	2.00	13.39	13.00	81.00	74.00	73.00	8.00	10.00
3.00	18.00	2.50	13.45	24.00	72.00	20.00	33.00	8.50	10.00
2.00	18.00	2.50	13.63	46.00	55.00	91.00	91.00	2.10	2.50
9.00	13.00	3.00	13.59	50.00	55.00	74.00	80.00	2.40	3.00
1.00	19.00	3.00	13.69	20.00	71.00	80.00	86.00	2.50	3.00
4.00	17.00			28.00	72.00			2.60	3.00
7.00	15.00			45.00	56.00			2.70	3.00
9.00	13.00			48.00	59.00			3.00	4.00
2.00	19.00							3.20	4.00
1.00	20.00							3.40	4.00
8.00	14.00							3.80	5.00
6.00	15.00							4.10	5.00
6.00	15.00							4.50	6.00
9.00	13.00							4.90	6.00
7.00	15.00							5.00	6.00
5.00	17.00							5.20	6.00
4.00	17.00							5.50	7.00
3.00	18.00							5.80	7.00
7.00	15.00							6.40	7.00
7.00	15.00							6.80	8.00
1.00	20.00							6.90	8.00
9.00	13.00							7.00	10.00
8.00	15.00							9.00	10.00

22		13							
x	**y**	**x**	**y**	**x**	**y**	**x**	**y**	**x**	**y**
15.00	95.00	44	45	65	61	80	74	79	73
26.00	71.00	75	68	67	65	79	73	80	75
10.00	83.00	75	68	68	65	43	41	67	62
9.00	91.00	80	74	78	73	75	65	47	43
15.00	102.00	79	73	80	76	77	71	70	65
20.00	87.00	75	67	81	74	72	65	80	75
18.00	93.00	74	71	82	76	76	70	76	67
11.00	100.00	53	53	74	70	68	63	81	76
8.00	101.00	78	73	55	51	45	44	59	58
20.00	91.00	76	66	78	73	44	45	65	61
7.00	113.00	79	73	75	64	75	68	67	65
9.00	96.00	64	59	71	67	75	68	68	65
10.00	83.00	78	72	57	54	80	74	78	73
11.00	84.00	66	60	65	62	79	73	80	76
11.00	102.00	67	57	77	68	75	67	81	74
10.00	100.00	75	69	72	66	74	71	82	76
12.00	105.00	50	47	77	69	53	53	74	70
42.00	57.00	50	46	70	66	78	73	55	51
17.00	121.00	52	50	73	67	76	66	78	73
11.00	86.00	58	55	81	75	79	73	75	64
10.00	100.00	81	74	80	73	64	59	71	67
		44	41	67	61	78	72	57	54
		78	71	57	54	66	60	65	62
		69	67	81	74	67	57	77	68
		75	69	70	66	75	69	72	66
		79	76	58	57	50	47	77	69
		77	70	78	71	50	46	70	66
		78	74	75	72	52	50	73	67
		77	69	67	63	58	55	81	75
		79	73	68	63	81	74	80	73
		70	66	77	69	44	41	67	61
		73	67	78	71	78	71	57	54
		63	60	75	69	69	67	81	74
		69	64	74	64	75	69	70	66
		76	67	46	43	79	76	58	57
		54	51	74	68	77	70	78	71
		80	72	70	66	78	74	75	72
		82	74	58	55	77	69	67	63
		58	52	79	73	79	73	68	63
		52	48	55	54	70	66	77	69
		76	69	68	62	73	67	78	71
		79	73	81	74	63	60	75	69
		80	75	81	75	69	64	74	64
		67	62	82	75	76	67	46	43
		47	43	68	65	54	51	74	68
		70	65	78	72	80	72	70	66
		80	75	45	41	82	74	58	55
		76	67	72	65	58	52	79	73
		81	76	73	69	52	48	55	54
		59	58	74	70	76	69	68	62